The Deviation Matrix of Continuous-Time Markov Processes

Leder, N.

2010

document version
Publisher's PDF, also known as Version of record

Chapter 1

The Deviation Matrix

1.1 Preliminaries

For some set of index parameters $\mathcal{T}$, let $\mathcal{X} \stackrel{\text{def}}{=} \{X_t,\ t \in \mathcal{T}\}$ be a family of random variables each defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ where Ω is an abstract set of individual elements ω which is supposed to be non-empty throughout this thesis. Then $\mathcal{F}$ is some σ-field of subsets of Ω and $\mathcal{P}$ denotes the corresponding probability measure which associates with every set in $\mathcal{F}$ the probability that it takes place. Each random variable is a mapping

$$X_t :\ (\Omega, \mathcal{F}, \mathcal{P}) \to (S, \mathcal{S}), \qquad \omega \mapsto X_t(\omega),$$

where $(S, \mathcal{S})$ is a measurable space consisting of state space S which is assumed to be denumerable as long as not demanded differently, and $\mathcal{S}$ the corresponding σ-field and it holds that

$$\{\omega \in \Omega :\ X_t(\omega) \in S'\} \in \mathcal{F}, \qquad \forall S' \in \mathcal{S}.$$

Finally, the parameter set $\mathcal{T}$ reflects the time scale and since this thesis is devoted to continuous-time processes we set $\mathcal{T} = \mathbb{R}_+$. Whenever we intend to highlight the connection to discrete-time processes $\mathcal{T}$ becomes $\mathbb{N}_0$ what we will reflect in our notation by replacing X_t by X_n within the respective formulas. To avoid intricate formulas in later computations we shorten notation and replace $\{\omega \in \Omega :\ X_t(\omega) \in S'\}$ by

$$\{X_t \in S'\}, \qquad t \geq 0.$$

A corresponding representation is $\{X_t = x\}$ for $x \in S$.

The existence of a stochastic process $\mathcal{X}$ on appropriate $(\Omega, \mathcal{F}, \mathcal{P})$ is assured by Kolmogorov's theorem and the fact that every denumerable state space S is a Polish space. Provided the left-hand side is defined, a stochastic process satisfying

$$\mathcal{P}\{X_t = y | X_{t_1} = x_1, X_{t_2} = x_2, \ldots, X_{t_n} = x_n\} = \mathcal{P}\{X_t = y | X_{t_n} = x_n\}, \\ \forall x_1, \ldots, x_n, y \in S, \ \forall t > t_n > \cdots > t_1 \geq 0,$$

that is, its future development does not depend on the whole history but only on the current state, is called a *Markov process*. All Markovian processes appearing in this thesis have the following property

$$\mathcal{P}\{X_{t+s} = y | X_s = x\} = \mathcal{P}\{X_t = y | X_0 = x\}, \qquad x, y \in S, \ 0 \leq s, t,$$

reflected by naming them time-homogeneous and are uniquely defined by their *transition probability matrix* $P(t) \stackrel{\text{def}}{=} (p(x, y; t))_{x,y \in S}$ with elements

$$p(x, y; t) \stackrel{\text{def}}{=} \mathcal{P}\{X_t = y | X_0 = x\}, \qquad x, y \in S.$$

Throughout this thesis, we will require $P(t)$ to be *standard*, i.e., it holds that

$$\lim_{t \to 0} p(x, y; t) = p(x, y; 0) = \delta(x, y), \qquad x, y \in S,$$

where $\delta(x, y)$ denotes Kronecker's delta, and in matrix notation $\lim_{t \to 0} P(t) = I$ with I the identity matrix of appropriate size and we define the *limiting matrix*

$$\Pi \stackrel{\text{def}}{=} \lim_{t \to \infty} P(t)$$

with rows equal to the limiting distributions

$$\pi(x, y) \stackrel{\text{def}}{=} \lim_{t \to \infty} p(x, y; t), \qquad x, y \in S.$$

The limiting matrix is often called the *ergodic projector* as for any given initial distribution, say μ, $\mu\Pi$ yields the process' asymptotic distribution. In case that the process $\mathcal{X}$ is *positive recurrent*, implying that the expected first return time for any state is finite, and *irreducible* so that from any state $x \in S$ every state $y \in S$ is reachable with positive probability, it holds for the entries of the limiting matrix that

$$\pi(x, y) = \pi(z, y) > 0, \qquad \forall x, y, z \in S,$$

i.e., all rows are equal and the entries strictly positive, what we reflect by using short-hand notation $\pi(y) = \pi(x, y)$, $x, y \in S$. For such a process the *stationary distribution* which is any probability measure satisfying

$$\gamma(y) = \sum_{x \in S} \gamma(x) p(x, y; t), \qquad x, y \in S,$$

is uniquely determined and equal to the limiting distribution. If we relax the assumption of irreducibility by separating state space S into several subsets S_i within which irreducibility holds while

$$p(x, y; t) = 0, \qquad x \in S_{i_1},\ y \in S_{i_2}, \qquad S_{i_1} \neq S_{i_2},$$

we arrive at a so-called *multi-chain* $\mathcal{X}$ consisting of several *ergodic* classes S_i so that all $\mathcal{X}_i$ defined on S_i are positive recurrent and irreducible. By arranging rows and columns in proper order such a process has transition probability and limiting matrix

$$P(t) = \begin{pmatrix} P_1(t) & 0 & 0 & \cdots \\ 0 & P_2(t) & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \quad \text{and} \quad \Pi = \begin{pmatrix} \Pi_1 & 0 & 0 & \cdots \\ 0 & \Pi_2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

where the 0s are matrices of appropriate size having zero-entries only, that is, the matrices $P(t)$ and Π show block structure (a property which carries over to the later on introduced generator and deviation matrix associated with these kind of processes) with each Π_i having identical rows. Such a process $\mathcal{X}$ has no unique stationary distribution but one which is dependent on the initial distribution μ with which the process starts at time 0, i.e., the distribution of X_0.

In case that state space S has exactly one ergodic class we call $\mathcal{X}$ a *uni-chain*. Please note that for any Markov process having at least one ergodic class it holds for the entries of the limiting matrix that

$$\sum_{y \in S} \pi(x, y) = 1, \qquad \forall x \in S.$$

In what follows we will use the expressions positive recurrence, irreducibility and ergodicity for processes and states alike.

Another kind of states are the *transient* ones to which the process does not necessarily return within finite time and which are not reachable from any ergodic state. Let $\mathcal{X}$ on S be a continuous-time Markov process having one ergodic class and some transient states. By separating state space $S = S_t \cup S_e$ where the subscripts t and e reflect transient and ergodic states, respectively, we get for the transition probability and limiting matrix

$$P(t) = \begin{pmatrix} P_{ee}(t) & 0 \\ P_{te}(t) & P_{tt}(t) \end{pmatrix} \qquad \text{and} \qquad \Pi = \begin{pmatrix} \Pi_{ee} & 0 \\ \Pi_{te} & 0 \end{pmatrix} \tag{1.1}$$

where the 0s are matrices of appropriate size having zero-entries only and Π_{ee} as well as Π_{te} are matrices with all rows equal to the limiting distribution of the ergodic class π_e.
The matrix $Q \stackrel{\text{def}}{=} (q(x,y))_{x,y \in S}$ with intensities

$$\begin{aligned} q(x,y) &\stackrel{\text{def}}{=} \lim_{t \to 0} \frac{\mathcal{P}\{X_t = y | X_0 = x\}}{t}, \qquad x \neq y, \ x, y \in S, \\ q(x,x) &\stackrel{\text{def}}{=} -\lim_{t \to 0} \frac{1 - \mathcal{P}\{X_t = x | X_0 = x\}}{t}, \qquad x \in S, \end{aligned}$$

is called the *infinitesimal generator* of $\mathcal{X}$ (also known as Q-matrix). For all processes appearing in this thesis Q is *conservative*, that is, it satisfies

$$q(x) \stackrel{\text{def}}{=} -q(x,x) = \sum_{x \neq y} q(x,y), \qquad x \in S.$$

In case that additionally it holds that

$$\sup_{x \in S} q(x) < \infty \tag{1.2}$$

we call the process $\mathcal{X}$ *uniformizable.*

Remark 1.1.1 *In case that $\mathcal{X}$ is uniformizable and its generator Q is conservative the associated transition probability matrix $P(t)$ satisfies Kolmogorov's backward equation*

$$\frac{dP(t)}{dt} = QP(t)$$

and it holds that

$$P(t) = \exp[Qt], \qquad t \geq 0,$$

(see, e.g., Theorem 4.18 in [80], p.196).

In the following lemma we will summarize some properties of the matrices associated with $\mathcal{X}$.

Lemma 1.1.2 *Let $\mathcal{X}$ be a continuous-time Markov process with transition matrix $P(t)$, $t \geq 0$, infinitesimal generator Q and limiting matrix Π then it holds*

(i) $\Pi^2 = \Pi$,

(ii) $\Pi Q = Q\Pi = \mathbf{0}$ *and* $A\Pi = \mathbf{0}$,

(iii) $\Pi P(t) = P(t)\Pi = \Pi$ *and* $B\Pi = \Pi$,

where $A, B \in \mathbb{R}^{S\times S}$ are matrices with $A_{(x,y)} = B_{(x,y)} = 0$ whenever x and y are of different ergodic classes and

$$\sum_{y\in S} A_{(x,y)} = 0 \qquad \text{and} \qquad \sum_{y\in S} B_{(x,y)} = 1, \qquad \forall x \in S.$$

Furthermore $\mathbf{0}$ *is a matrix of appropriate size with all elements equal to zero.*

This lemma was proven in [32] for the uni-chain case. However, the extension to multi-chains and processes having some transient states follows easily from the block structure of the associated matrices. We now introduce the protagonist of this thesis:

Definition 1.1.3 (Deviation Matrix) *Let $\mathcal{X}$ be a continuous-time Markov process with transition matrix $P(t)$, infinitesimal generator Q and limiting matrix Π then we define $D \stackrel{\text{def}}{=} (d(x,y))_{x,y\in S}$ by*

$$d(x,y) \stackrel{\text{def}}{=} \int_0^\infty (p(x,y;t) - \pi(x,y))\, dt, \qquad x, y \in S,^{1} \tag{1.3}$$

which is said to exist whenever all integrals are finite and in that case D is called the deviation matrix *of $\mathcal{X}$.*

Note that another common expression for the deviation matrix is the *fundamental matrix* (see, e.g., [76, 80]), but since fundamental matrix is also used for $F \stackrel{\text{def}}{=} (\Pi - Q)^{-1}$ (see [32]), we will stick to the term deviation matrix.
The following lemma presents the main properties of the deviation matrix.

[1] In what follows and whenever they are not required, we will do without brackets in integrals to simplify formulas.

Lemma 1.1.4 *Let $\mathcal{X}$ be a continuous-time Markov process with transition matrix $P(t)$, infinitesimal generator Q, limiting matrix Π and deviation matrix D then it holds*

(i) $D\mathbf{1} = \mathbf{0}$ *with* $\mathbf{1}$ *an appropriately sized vector with all entries equal to one,*

(ii) $D\Pi = \Pi D = \mathbf{0}$*,*

(iii) $DQ = QD = \Pi - I$ *which implies that* D *solves the Poisson equation.*

The proof for this lemma was provided in [32] for a uni-chain Markov process. But analogously to Lemma 1.1.2, we can easily extend their findings to multi-chains by recalling the fact that all matrices consist of blocks having uni-chain properties or zero entries. Therefore, we will focus now on the less obvious case of a process having some transient states. With (1.1) and

$$D = \begin{pmatrix} D_{ee} & 0 \\ D_{te} & D_{tt} \end{pmatrix} = \begin{pmatrix} \int_0^\infty P_{ee}(t) - \Pi_{ee} dt & 0 \\ \int_0^\infty P_{te}(t) - \Pi_{te} dt & \int_0^\infty P_{tt}(t) dt \end{pmatrix}$$

property (i) follows from the fact that $P(t)$ and Π have row sums equal to 1 (interchange of integration and summation formally justified in [32]) and (ii) follows with Lemma 1.1.2. Property (iii) holds for transient processes as well because it follows directly with the Kolmogorov forward and backward equation which is not restricted to ergodic Markov processes but can be extended to transient chains (see, e.g., [80], p.245).

The main tool for our analysis is the *weighted supremum norm*, also called v-norm, denoted by $\|\cdot\|_v$, where v is some vector with elements $v(x) \geq 1$ for all $x \in S$, and for any $w \in \mathbb{R}^S$

$$||w||_v \stackrel{\text{def}}{=} \sup_{x \in S} \frac{|w(x)|}{v(x)}.$$

For a matrix $A \in \mathbb{R}^{S \times S}$ the v-norm is given by

$$\|A\|_v \stackrel{\text{def}}{=} \sup_{x, ||w||_v \leq 1} \frac{\sum_{y=1}^S |A(x,y) w(y)|}{v(x)}. \tag{1.4}$$

1.2 The Discrete-time Analogon

As the deviation matrix is predominantly discussed for discrete-time Markov chains in the literature we will present within this section the discrete-time analogon of (1.3) and show the connection between the two different time scales.
Let $\tilde{\mathcal{X}} \stackrel{\text{def}}{=} \{\tilde{X}_n,\ n \in \mathbb{N}_0\}$ be a discrete-time Markov chain on state space S with an n-step transition probability matrix $\tilde{P}(n) \stackrel{\text{def}}{=} (\tilde{p}(x, y; n))_{x,y\in S}$ having entries

$$\tilde{p}(x, y; n) \stackrel{\text{def}}{=} \mathcal{P}\{\tilde{X}_n = y | \tilde{X}_0 = x\}, \qquad x, y \in S,$$

which satisfies

$$\tilde{P}(n) = \tilde{P}^n, \qquad n \geq 0,$$

where $\tilde{P} \stackrel{\text{def}}{=} \tilde{P}(1)$. Suppose that $\tilde{\mathcal{X}}$ is aperiodic, i.e., $gcd\{n \geq 1 : p(x, x; n) > 0\} = 1$. Then its limiting matrix $\tilde{\Pi} \stackrel{\text{def}}{=} (\tilde{\pi}(x, y))_{x,y\in S}$ has elements

$$\tilde{\pi}(x, y) \stackrel{\text{def}}{=} \lim_{n\to\infty} \tilde{p}(x, y; n), \qquad x, y \in S.$$

The next definition is the analogon to Definition 1.1.3.

Definition 1.2.1 (Discrete-time Deviation Matrix) *Let $\tilde{\mathcal{X}}$ be a discrete-time Markov chain with transition matrix $\tilde{P}$ and limiting matrix $\tilde{\Pi}$ then we define $\tilde{D} \stackrel{\text{def}}{=} (\tilde{d}(x, y))_{x,y\in S}$ with*

$$\tilde{d}(x, y) \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} \left(\tilde{p}(x, y; n) - \tilde{\pi}(x, y)\right), \qquad x, y \in S, \tag{1.5}$$

which is said to exist whenever all sums are finite and in that case $\tilde{D}$ is called the deviation matrix *of $\tilde{\mathcal{X}}$.*

The matrix representation of (1.5) is

$$\tilde{D} = \sum_{n=0}^{\infty} \left(\tilde{P}^n - \tilde{\Pi}\right). \tag{1.6}$$

Remark 1.2.2 *An alternative way to define the limiting matrix as well as the deviation matrix is via the respective Cesaro limit given by*

$$\tilde{\Pi} = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N} \tilde{P}^k \qquad \text{and} \qquad \tilde{D} = \lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N}\sum_{n=0}^{k} \left(\tilde{P}^n - \tilde{\Pi}\right)$$

where the existence of the previous one was stated and proven in Theorem 3.3.1 of [131], p.96f, and the associated proof whereas the latter one directly follows with our assumption of the existence of $\tilde{D}$. The advantage of this definition is that it overcomes the limitation to aperiodic Markov chains. However, since periodicity does not occur for homogeneous continuous-time Markov processes we will stick to (1.5).

Now we can use uniformization theory to provide the connection between continuous- and discrete-time Markov processes. Recall that in accordance with (1.2) all generator matrix entries of a uniformizable process are finite and we can identify an upper bound ν so that

$$q(x) < \nu < \infty, \qquad \forall x \in S.$$

Then we can introduce the transition probability matrix $\tilde{P}$ by

$$\tilde{P} \stackrel{\text{def}}{=} I + \frac{1}{\nu}Q. \tag{1.7}$$

The parameter ν is the intensity of a Poisson process $\{N(t),\ t \geq 0\}$ which counts the number of state changes that process $\mathcal{X}$ realizes within a time interval $[0,t)$. It provides the connection between a continuous-time process $\{X_t,\ t \in \mathbb{R}_+\}$ and a discrete one $\{\tilde{X}_n,\ n \in \mathbb{N}_0\}$. Replacing n by $N(t)$ transforms $\{\tilde{X}_n,\ n \in \mathbb{N}_0\}$ to the *subordinated chain* $\{\tilde{X}_{N(t)},\ t \geq 0\}$ which is governed by the same transition probability matrix $\tilde{P}$, but instead of formerly deterministic transition-instants has random jump-times. According to Theorem 4.19 in [80], p.198, a uniformizable Markov process $\mathcal{X}$ and its subordinated chain are equally distributed what we reflect by

$$\{X_t,\ t \geq 0\} \stackrel{d}{=} \{\tilde{X}_{N_t},\ t \geq 0\}$$

and it holds that

$$P(t) = \tilde{P}(N_t) = e^{-\nu t}\sum_{n=0}^{\infty} \tilde{P}^n \frac{(\nu t)^n}{n!}, \qquad t \geq 0,$$

and

$$\Pi = \tilde{\Pi}.$$

We introduce now the Laplace transformation of D

$$\hat{D}(\alpha) \stackrel{\text{def}}{=} \int_0^\infty e^{-\alpha t} \left(P(t) - \Pi\right) dt, \qquad \alpha > 0.$$

Remark 1.2.3 *Note that the Laplace transformation of the deviation matrix is closely related to another common tool in Markov theory: the Laplace transformation of the transition probability matrix better known as the* resolvent

$$\hat{P}(\alpha) \stackrel{\text{def}}{=} \int_0^\infty e^{-\alpha t} P(t) dt, \qquad \alpha > 0,$$

with the connection provided by

$$\hat{D}(\alpha) = \hat{P}(\alpha) - \frac{\Pi}{\alpha}, \qquad \alpha > 0.$$

By using its Laplace transformation instead of D we justify the interchange of summation and integration within the following transformation.

$$\begin{aligned}
\hat{D}(\alpha) &= \int_0^\infty e^{-\alpha t} \left(P(t) - \Pi\right) dt = \int_0^\infty e^{-\alpha t} \left(\tilde{P}(N_t) - \tilde{\Pi}\right) dt \\
&= \int_0^\infty e^{-(\alpha+\nu)t} \sum_{n=0}^\infty \left(\tilde{P}^n - \tilde{\Pi}\right) \frac{(\nu t)^n}{n!} dt \\
&= \sum_{n=0}^\infty \left(\tilde{P}^n - \tilde{\Pi}\right) \frac{\nu^n}{n!} \int_0^\infty t^n e^{-(\alpha+\nu)t} dt, \qquad \alpha > 0.
\end{aligned}$$

Since $\int_0^\infty t^n e^{-(\alpha+\nu)t} dt = n!(\alpha+\nu)^{-(n+1)}$ we arrive at

$$\sum_{n=0}^\infty \left(\tilde{P}^n - \tilde{\Pi}\right) \frac{\nu^n}{n!} \int_0^\infty t^n e^{-(\alpha+\nu)t} dt = \sum_{n=0}^\infty \left(\tilde{P}^n - \tilde{\Pi}\right) \frac{\nu^n}{(\alpha+\nu)^{n+1}}.$$

By a Tauberian theorem we get $\lim_{\alpha \to 0} \hat{D}(\alpha) = D$ (see, e.g., Theorem 3a in [136], p.186) and with Abel's theorem (see, e.g., Theorem B.2 in [80], p.304) it holds that

$$\lim_{\alpha \to 0} \sum_{n=0}^\infty \left(\tilde{P}^n - \tilde{\Pi}\right) \frac{\nu^n}{(\alpha+\nu)^{n+1}} = \frac{1}{\nu} \tilde{D}$$

what finally yields the connecting equation

$$D = \frac{1}{\nu}\tilde{D}.$$

Remark 1.2.4 *Based on (1.7) it is always possible to construct a continuous-time process $\mathcal{X}$ associated with the discrete-time chain $\tilde{\mathcal{X}}$. Contrarily, it is not necessarily possible to construct a discrete-time version of a continuous-time Markov process via uniformization because this procedure requires that $\sup_{x\in S} q(x) < \infty$, a property which, e.g., the queue-length process of the $M/M/\infty$ system does not satisfy. To allow for a wide range of processes, the formulas and derivations presented in this thesis are devoted to continuous-time processes which need not be uniformizable. All findings are transferable to a discrete-time setting via uniformization theory.*

1.3 Existence

In Definition 1.1.3 we stated that the deviation matrix exists whenever all integrals in (1.3) are finite. Whenever we have a closed-form representation of D we can surely check the finiteness of its elements directly. However, we will show in later chapters even closed-form representations can be of intricate form (see, e.g., the deviation matrix of an inventory system in Chapter 5) for which finiteness cannot be seen on first sight. Therefore, we provide two sufficient conditions assuring $d(x, y) < \infty$ for all $x, y \in S$.
From (1.4) we know that a finite v-norm of D with the additional conditions $v(x) < \infty$ for all $x \in S$ and $\inf_{x\in S} v(x) > 0$ implies

$$\sup_{y\in S} |d(x, y)| \leq \|D\|_v v(x) < \infty, \qquad \forall x \in S,$$

so that we already arrive at the first sufficient condition

(C1)

$$\|D\|_v < \infty$$

for some finite v-norm.

Note that v-geometric ergodicity of the process $\mathcal{X}$, i.e., there exists some finite constant ρ and $\delta \in (0, 1)$ such that

$$\|P(t) - \Pi\|_v \leq \rho\delta^t, \qquad \forall t \geq 0,$$

implies that

$$\|D\|_v \leq \int_0^\infty \|P(s) - \Pi\|_v ds \leq \rho \int_0^\infty \delta^s ds = -\frac{\rho}{ln[\delta]} < \infty$$

(for sufficient conditions of v-geometric ergodicity we refer, e.g., to [65, 112]). The advantage of condition **(C1)** is that we have to check a single value only instead of assuring the finiteness for every element of D. However, **(C1)** requires the identification of an appropriate norm since $\|D\|_v = \infty$ does not necessarily imply that there exist some $x, y \in S$ with $d(x, y) = \infty$ but $\|D\|_v = \infty$ could also be due to the choice of an inappropriate v-norm. In case of finite S all norms are equivalent so that the most advisable choice is $v(x) = 1$ for all $x \in S$. But if S becomes denumerable or even general the identification of an optimal $v(x)$ is essential and no longer obvious.
Therefore we provide an alternative approach presented by Coolen-Schrijner and van Doorn in [32] for which they require $\mathcal{X}$ to be irreducible and positive recurrent. On the basis of mean first passage times from state x to state y, $x, y \in S$, denoted by $m(x, y)$ they state in their Theorem 4.1 that the finiteness of all elements of D is assured by

$$m_\pi(y) \stackrel{\text{def}}{=} \sum_{x \in S} \pi(x) m(x, y) < \infty$$

for at least one $y \in S$ because then it follows by their Theorem 3.1 that $m_\pi(y) < \infty$ for all $y \in S$. The connection between mean first passage times and the deviation matrix as well as the fact that finiteness of the former ones causes finiteness of the elements of the latter one becomes apparent by the common representation

$$d(x, y) = \pi(y) \left(m_\pi(y) - m(x, y) \right), \qquad x, y \in S,$$

(see (5.8) in [32] and (1.35) herein), which is presented later in this thesis.
This leads to the next sufficient condition for the existence of the deviation matrix.

(C2)

$$m_\pi(y) < \infty$$

for at least one $y \in S$.

The obvious advantage is that we no longer have to determine an optimal v-norm. However, now we have to check the finiteness of the mean first passage time what becomes tedious if the closed-form representation of $m_\pi(y)$ is intricate.

1.4 Formulas

While our definition of the deviation matrix in (1.3) provides for an insight into its properties and hints at the applications for which it might be useful, it is not a good formula for the computation of definite examples like the deviation matrices of queueing systems. Therefore, we will look for alternative representations which overcome this restriction.

For the sake of completeness we will first of all summarize some closed-form expressions which can already be found in the literature before we present alternative formulas in Section 1.4.1, 1.4.2, 1.4.3 and 1.4.4. First of all the deviation matrix of a single server M/M/1 queue with arrival rate $\lambda > 0$, service rate $\mu > 0$ and traffic rate $\rho \stackrel{\text{def}}{=} \frac{\lambda}{\mu}$

$$d^{M/M/1}(x,y) = \frac{\rho^{\max\{y-x,0\}} - (x+y+1)(1-\rho)\rho^y}{\mu(1-\rho)}, \qquad x,y \in \{0,1,\dots\},$$

which exists whenever $\rho < 1$, and an M/M/1/N queue having the same rates but without the restriction on ρ

$$d^{M/M/1/N}(x,y) = \begin{cases} -\frac{\max\{x-y,0\}}{\mu} + \frac{3x(x+1)+3(N-y)(N+1-y)-N(N+2)}{6\mu(N+1)}, & \lambda = \mu, \\ \frac{\rho^{\max\{y-x,0\}}}{\mu(1-\rho)} + \frac{\rho^y\left(\rho^{N-x+1}+\rho^{N-y+1}-(x+y-1)(1-\rho)-2\right)}{\mu(1-\rho)(1-\rho^{N+1})} & \lambda \neq \mu, \\ +\frac{2\rho^{y+1}\left(1-(N+1)\rho^N+N\rho^{N+1}\right)}{\mu(1-\rho)(1-\rho^{N+1})^2} \end{cases} \tag{1.8}$$

were provided in [83] while the more general case with $\rho = \frac{\lambda}{c\mu}$

$$\begin{aligned} d^{M/M/c/N}(x,y) &= \frac{(\pi(c))^2\left(\rho-(1-\rho)(2N-2c+1)\rho^{N+1-c}-\rho^{2N-2c+2}\right)}{c\mu(1-\rho)^3} \\ &\quad +\frac{c!\rho^{-c}(\pi(c))^2}{c^c\lambda}\sum_{z<c}\frac{z!}{\rho^z c^z}\left(\sum_{v\le z}\frac{\rho^v c^v}{v!}\right)^2, \qquad x=y=c, \end{aligned}$$

and for $x, y \neq c$

$$
\begin{aligned}
&d^{M/M/c/N}(x,y)\\
&\quad= \frac{\rho^{y-c} c!\, d(c,c)}{\min\{y,c\}! c^{c-\min\{y,c\}}} + \frac{\rho^{\max\{\max\{y,c\}-\max\{x,c\},0\}} - \rho^{\max\{y,c\}-c}}{c\mu(1-\rho)}\\
&\qquad+ \frac{\pi(c)\, c! \rho^{y-c} \left(2c - \max\{x,c\} - \max\{y,c\}\right)}{\min\{y,c\}! c^{c-\min\{y,c\}} c\mu(1-\rho)}\\
&\qquad+ \frac{\pi(c)\, c! \rho^{y-c}}{\min\{y,c\}! c^{c-\min\{y,c\}} c\mu(1-\rho)^2}\\
&\qquad\cdot \left(\rho^{N+1-c} + \rho^{N+1+y-\max\{x,c\}-c} - \rho^{N+1+y-2c} - \rho^{N+1+\max\{y,c\}-2c}\right)\\
&\qquad+ \frac{\rho^y c^y}{y!\mu} \left(\sum_{z=\min\{y,c\}+1}^{c} \frac{(z-1)!}{\rho^z c^z} - \sum_{z=\min\{y,c\}+1}^{\min\{x,c\}} \frac{(z-1)!}{\rho^z c^z} \right)\\
&\qquad- \frac{\pi(c)\, c! \rho^{y-c}}{\min\{y,c\}! c^{c-\min\{y,c\}} \mu} \left(\sum_{z=\min\{y,c\}+1}^{c} \sum_{v<z} \frac{(z-1)! \rho^{v-z}}{v! c^{z-v}} \right.\\
&\qquad\qquad\qquad\qquad \left. + \sum_{z=\min\{x,c\}+1}^{c} \sum_{v<z} \frac{(z-1)! \rho^{v-z}}{v! c^{z-v}} \right)
\end{aligned}
$$

where

$$
\pi(c) = \frac{1-\rho}{\sum_{z=1}^{c} \rho^z \frac{c! z}{(c-z)! c^{z+1}} - \rho^{N+1-c}}
$$

was presented in [86]. From that $D^{M/M/c}$ can be obtained by letting N tend to ∞.

Additional closed-form formulas for the deviation matrix of more exotic Markov processes were provided in [81] where Kirkland derived several examples of the *group generalized inverse* which is another expression for the deviation matrix and stems from its property

$$
D(-Q)D = D
$$

identifying it as a group generalized inverse of $-Q$. Starting with the transition probability matrix $\tilde{P}$ (in what follows we use (1.7) to transfer $\tilde{P}$ to our continuous-time setting) he gives the respective deviation matrices. However, his models are rather exotic and not necessarily interpretable.

$$
Q = \frac{1}{n-1} J - \frac{n}{n-1} I \qquad \Longrightarrow \qquad D = \frac{n-1}{n} I - \frac{n-1}{n^2} J
$$

where I is the $n \times n$-identity matrix and J an $n \times n$-matrix with all elements equal to one, and

$$Q = \begin{pmatrix} -I & \mathbf{1} \\ \frac{1}{n-1}\mathbf{1}^T & -1 \end{pmatrix} \qquad \Longrightarrow \qquad D = \begin{pmatrix} I - \frac{3}{4n-4}J & -\frac{1}{4}\mathbf{1} \\ -\frac{1}{4n-4}\mathbf{1}^T & \frac{1}{4} \end{pmatrix}$$

where $\mathbf{1}$ is an $n \times 1$-vector with all entries equal to one and $\mathbf{1}^T$ denotes its transpose.
After providing this literature survey we will proceed now with some representations of the deviation matrix which are especially useful to compute D. We will highlight each of the formulas by an example.

1.4.1 Matrix Inversion

Provided that all matrices appearing in the following formulas exist, we get by adding $\Pi D (= 0)$ to Lemma 1.1.4 (iii)

$$(\Pi - Q)D = I - \Pi. \tag{1.9}$$

With the sufficient condition $\|(I - \Pi + Q)^n\|_v < 1$ for some $n \in \mathbb{N}$ and an appropriately chosen v-norm we assure the existence of the inverse of $(\Pi - Q)$ so that we can solve (1.9) for D and get

$$D = (\Pi - Q)^{-1} - (\Pi - Q)^{-1}\Pi$$

and since $(I - \Pi + Q)^n \Pi = 0$ for $n > 0$ (see Lemma 1.1.2) implies

$$(\Pi - Q)^{-1}\Pi = \sum_{n=0}^{\infty}(I - \Pi + Q)^n \Pi = \Pi$$

we arrive at a simple formula for the deviation matrix

$$D = (\Pi - Q)^{-1} - \Pi \tag{1.10}$$

which is especially handy in case of finite, small state spaces S. If Q and Π are given for a process $\mathcal{X}$ the deviation matrix can be computed by a simple matrix inversion.

Example 1.4.1 (Deviation Matrix of Gordon-Newell Networks) *An example of finite state space Markov processes are Gordon-Newell networks introduced by Gordon and Newell in [49]. They presented a model with N* *single-type customers inside a network consisting of* J *nodes each node having unbounded waiting room and exponentially distributed service with parameter* $\mu_j > 0$ *at the* j^{th} *node. The associated process* $\mathcal{X}$ *describes customers' spreading over the* J *nodes and the state space is*

$$S = \left\{ x = (x_1, \dots, x_J) \in \mathbb{N}_0^J,\ \sum_{i=1}^{J} x_i = N \right\}.$$

The probability to go to node j *after being served in* i *known as* routing probability *is denoted by* $r_{ij} \geq 0$ *for* $i, j \in \{1, \dots, J\}$ *and it holds*

$$\sum_{j=1}^{J} r_{ij} = 1 \qquad \text{and we assume} \qquad r_{ii} = 0, \qquad \forall i \in \{1, \dots, J\}.$$

Process $\mathcal{X}$ *has infinitesimal generator*

$$q(x, y) = \begin{cases} \min\{x_i, 1\}\mu_i r_{ij} & y = x - e_i + e_j,\ i \neq j, \\ -\sum_{i=1}^{J} \min\{x_i, 1\}\mu_i & y = x, \\ 0 & \text{otherwise}, \end{cases}$$

where e_i *denotes the* i^{th} *unit vector,* $i \in \{1, \dots, J\}$*. Then the stationary distribution is given by*

$$\pi(x) = K^{-1} \prod_{i=1}^{J} \frac{\lambda_i^{x_i}}{\mu_i^{x_i}}, \qquad x \in S, \tag{1.11}$$

where

$$K = \sum_{x \in S} \prod_{i=1}^{J} \frac{\lambda_i^{x_i}}{\mu_i^{x_i}}$$

and λ_i *is any solution of*

$$\lambda_j = \sum_{i=1}^{J} \lambda_i r_{ij}.$$

Let, e.g., $\mathcal{X}$ *be* N *customers' spreading over a cyclic network of* J *nodes (see Figure 1.1) each providing exponentially-*μ *distributed service. Such a system has routing probabilities*

$$r_{ij} = \begin{cases} 1 & i < J,\ j = i + 1, \\ 1 & i = J,\ j = 1, \\ 0 & \text{otherwise}, \end{cases}$$

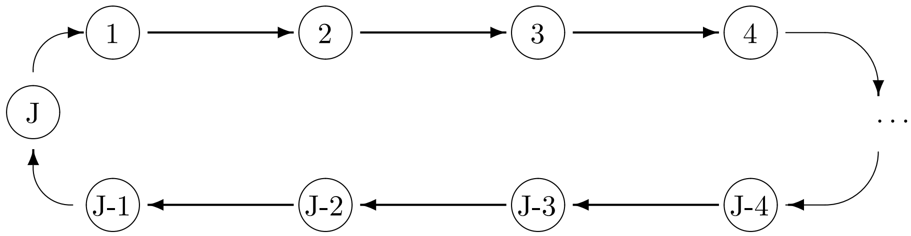

Figure 1.1: Cyclic Network

and a Q-matrix with entries

$$q(x,y) = \begin{cases} \min\{x_i, 1\}\mu & y = x - e_i + e_j,\ i \neq j, \\ -\mu \sum_{i=1}^{J} \min\{x_i, 1\} & y = x, \\ 0 & \text{otherwise,} \end{cases}$$

and it follows with (1.11) that

$$\pi(x) = \binom{N+J-1}{N}^{-1}$$

where $|S| \stackrel{\text{def}}{=} \binom{N+J-1}{N}$ *is the cardinality of state space* S. *For* $N = 2$ *and* $J = 3$ *we have* $|S| = 6$ *so that* $\pi(x) = \frac{1}{6}$ *for all* $x \in S$ *and inserting these values into (1.10) yields the deviation matrix* D

$$\frac{1}{36\mu} \begin{pmatrix} 23 & -1 & -13 & 5 & -7 & -7 \\ -13 & 23 & -1 & -7 & 5 & -7 \\ -1 & -13 & 23 & -7 & -7 & 5 \\ -7 & 5 & -7 & 11 & -1 & -1 \\ -7 & -7 & 5 & -1 & 11 & -1 \\ 5 & -7 & -7 & -1 & -1 & 11 \end{pmatrix}$$

and for $N = 3$ *and* $J = 3$ *we have* $\pi(x) = \frac{1}{10}$ *and for the deviation matrix* D

$$\frac{1}{2600\mu} \begin{pmatrix} 2261 & -139 & -939 & 851 & -349 & -749 & -79 & 121 & -679 & -299 \\ -939 & 2261 & -139 & -749 & 851 & -349 & -679 & -79 & 121 & -299 \\ -139 & -939 & 2261 & -349 & -749 & 851 & 121 & -679 & -79 & -299 \\ -79 & 121 & -679 & 1111 & -89 & -489 & 181 & 381 & -419 & -39 \\ -679 & -79 & 121 & -489 & 1111 & -89 & -419 & 181 & 381 & -39 \\ 121 & -679 & -79 & -89 & -489 & 1111 & 381 & -419 & 181 & -39 \\ 851 & -349 & -749 & 341 & -259 & -459 & 1111 & -89 & -489 & 91 \\ -749 & 851 & -349 & -459 & 341 & -259 & -489 & 1111 & -89 & 91 \\ -349 & -749 & 851 & -259 & -459 & 341 & -89 & -489 & 1111 & 91 \\ -299 & -299 & -299 & 91 & 91 & 91 & -39 & -39 & -39 & 741 \end{pmatrix}.$$

1.4.2 State Reduction

In the previous Section, we presented a formula for the deviation matrix which requires matrix inversion a task which becomes computationally complex and numerically instable whenever the state space S is large and the matrix $\Pi - Q$ has values close to zero. An alternative approach was presented in [62, 125, 126] where Heyman and O'Leary, Sheskin as well as Sonin replaced the inversion by a recursive algorithm based on state space reduction. While these papers are devoted to discrete-time Markov chains we transfer their findings to the continuous-time setting.
Let $\mathcal{X}$ be a uni-chain continuous-time Markov process on finite state space S. The next lemma states an important precondition which holds for the deviation matrix of such a process.

Lemma 1.4.2 *In case of a finite state space S, the deviation matrix D for given Q and Π is uniquely determined by the right-hand equations of (ii) and (iii) in Lemma 1.1.4.*

Proof: Suppose we have a second matrix D^* solving the right-hand side of Lemma 1.1.4 (iii), i.e., $QD^* = \Pi - I$, then it holds

$$Q(D - D^*) = \mathbf{0}$$

where $\mathbf{0}$ denotes the $n \times n$-matrix with all entries equal to zero. According to Theorem 4.16 (1) in [15], p.156, Q is rank-deficient with rank $n-1$, and it is furthermore obvious that $\mathbf{0}$ has rank 0. Now we can apply Sylvester's rank inequality

$$\underbrace{\mathrm{rank}(Q)}_{=n-1} + \mathrm{rank}(D - D^*) - n \leq \underbrace{\mathrm{rank}\big(Q(D - D^*)\big)}_{=0}$$

from which we obtain that $D - D^*$ has to be of rank smaller or equal to 1. This implies that all rows of $D - D^*$ have to be multiples of a certain row-vector $x \in \mathbb{R}^{1\times n}$. So we have

$$D - D^* = \begin{pmatrix} d_1x_1 & d_1x_2 & \dots & d_1x_n \\ d_2x_1 & d_2x_2 & \dots & d_2x_n \\ \vdots & \vdots & \ddots & \vdots \\ d_nx_1 & d_nx_2 & \dots & d_nx_n \end{pmatrix}$$

with at least one factor d_y, $y \in S$ not equal to zero (the case $d_y = 0$ for all $y \in S$ can be considered by setting $x = 0$). Elementwise written, equation (ii) of Lemma 1.1.4, i.e., $\Pi D = \mathbf{0}$, gives us

$$\sum_{z \in S} \pi(z)(D - D^*)_{(z,y)} = 0 \qquad \Leftrightarrow \qquad x_y \sum_{z \in S} \pi(z) d_z = 0, \qquad y \in S.$$

Since in case of a uni-chain $\mathcal{X}$ $\sum_{z \in S} \pi(z) d_z > 0$ we get $x_y = 0$ for all $y \in S$. Hence, we have $D - D^* = 0$ and D is the unique solution solving (ii) and (iii).

■

Now set $S_1 = \{1, 2, \ldots, n\}$ with the associated process denoted by $\mathcal{X}_1$, the infinitesimal generator by Q_1 and the limiting matrix by Π_1. Then we can define a second process $\mathcal{X}_2$ on the reduced state space $S_2 = S_1 \setminus \{n\}$ by its transition rates satisfying

$$q_2(x, y) = q_1(x, y) - \frac{q_1(x, n) q_1(n, y)}{q_1(n, n)}, \qquad x, y \in S_2. \tag{1.12}$$

Equation (1.12) is the continuous-time analogon of the state reduced transition probabilities provided by Kemeny and Snell in [78]. For the limiting distribution of the state reduced process it holds

$$\pi_2(y) = \frac{\pi_1(y)}{1 - \pi_1(n)}, \qquad y \in S_2. \tag{1.13}$$

With Lemma 1.1.2 we get

$$\sum_{z=1}^{n} \pi_1(z) q_1(z, n) = 0 \qquad \Leftrightarrow \qquad \pi_1(n) = -\sum_{z=1}^{n-1} \pi_1(z) \frac{q_1(z, n)}{q_1(n, n)}.$$

In accordance with (1.13) we can replace $\pi_1(z)$ by $(1 - \pi_1(n))\pi_2(z)$ and arrive at

$$\begin{aligned} \pi_1(n) &= -\sum_{z=1}^{n-1} (1 - \pi_1(n)) \pi_2(z) \frac{q_1(z, n)}{q_1(n, n)} \\ \Leftrightarrow \qquad \frac{\pi_1(n)}{1 - \pi_1(n)} &= -\sum_{z=1}^{n-1} \pi_2(z) \frac{q_1(z, n)}{q_1(n, n)} \end{aligned}$$

what we solve for $\pi_1(n)$ and obtain

$$\pi_1(n) = \left(\left(-\sum_{z=1}^{n-1} \pi_2(z) \frac{q_1(z,n)}{q_1(n,n)} \right)^{-1} + 1 \right)^{-1} . \tag{1.14}$$

Remark 1.4.3 *Equations (1.13) and (1.14) provide for a possibility to compute the limiting distribution recursively. After having derived the Q-matrices up to Q_n using (1.12), we can successively compute the limiting matrices starting with $\Pi_n = (1)$ and by applying (1.14) for the computation of the last column's entries we can afterwards compute the remaining elements with (1.13). The actual formula for the rows of any Π_h, $h \in S_2$ is*

$$\pi_h = \begin{pmatrix} (1-\pi_h(n-h+1))\pi_{h+1}^T \\ \pi_h(n-h+1) \end{pmatrix}^T$$

where π_h and π_{h+1} are row-vectors in $\mathbb{R}^{S_h}$ and $\mathbb{R}^{S_{h+1}}$, respectively, and

$$\pi_h(n-h+1) = \left(\left(-\sum_{z=1}^{n-h} \pi_{h+1}(z) \frac{q_h(z,n-h+1)}{q_h(n-h+1,n-h+1)} \right)^{-1} + 1 \right)^{-1} .$$

Using these findings we can derive a recursive representation of the deviation matrix. Elementwise written Equation (iii) in Lemma 1.1.2 becomes

$$\sum_{z=1}^{n} d_1(x,z) q_1(z,y) = \pi_1(y) - \delta(x,y), \qquad x,y \in S_1, \tag{1.15}$$

and setting $y = n$ and $x < n$

$$d_1(x,n)q_1(n,n) + \sum_{z=1}^{n-1} d_1(x,z)q_1(z,n) = \pi_1(n). \tag{1.16}$$

Solving (1.16) for $d_1(x,n)$ yields

$$d_1(x,n) = \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_1(x,z) \frac{q_1(z,n)}{q_1(n,n)}. \tag{1.17}$$

By multiplying (1.17) with $q_1(n,y)$, $y<n$, we get

$$d_1(x,n)q_1(n,y) = \pi_1(n)\frac{q_1(n,y)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_1(x,z)\frac{q_1(z,n)q_1(n,y)}{q_1(n,n)}. \tag{1.18}$$

Inserting (1.18) into (1.15) yields for $x,y\in S_2$

$$\sum_{z=1}^{n-1} d_1(x,z)q_1(z,y) - \sum_{z=1}^{n-1} d_1(x,z)\frac{q_1(z,n)q_1(n,y)}{q_1(n,n)} = \pi_1(y) - \pi_1(n)\frac{q_1(n,y)}{q_1(n,n)} - \delta(x,y).$$

By applying (1.12) we can transfer this equation to

$$\sum_{z=1}^{n-1} d_1(x,z)q_2(z,y) = \pi_1(y) - \pi_1(n)\frac{q_1(n,y)}{q_1(n,n)} - \delta(x,y), \qquad x,y\in S_2. \tag{1.19}$$

Subtracting (1.19) from the state reduced version of (1.15) gives us

$$\begin{aligned} \sum_{z=1}^{n-1} d_2(x,z)q_2(z,y) - \sum_{z=1}^{n-1} d_1(x,z)q_2(z,y) &= \pi_2(y) - \pi_1(y) + \pi_1(n)\frac{q_1(n,y)}{q_1(n,n)} \\ \Leftrightarrow \qquad \sum_{z=1}^{n-1} \left(d_2(x,z) - d_1(x,z)\right) q_2(z,y) &= \pi_1(n)\pi_2(y) + \pi_1(n)\frac{q_1(n,y)}{q_1(n,n)}, \\ & \qquad\qquad x,y\in S_2, \end{aligned}$$

where we applied for the right-hand side the previously introduced equation $\pi_2(y) = \frac{\pi_1(y)}{1-\pi_1(n)}$. By writing the equation in matrix-form we get

$$\left(D_2 - D_1^{(1)}\right) Q_2 = \pi_1(n)(\Pi_2 + A) \tag{1.20}$$

where $D_1^{(1)}$ is the upper left $(n-1)\times(n-1)$-matrix of D_1 and A is given by

$$A = \frac{1}{q_1(n,n)} \begin{pmatrix} q_1(n,1) & q_1(n,2) & \cdots & q_1(n,n-1) \\ q_1(n,1) & q_1(n,2) & \cdots & q_1(n,n-1) \\ \vdots & \vdots & \ddots & \vdots \\ q_1(n,1) & q_1(n,2) & \cdots & q_1(n,n-1) \end{pmatrix}.$$

Since $\sum_{z=1}^{n-1} q_1(n,z) = -q_1(n,n)$ the matrix $-A$ has constant row sums equal to 1. Transforming (1.20) gives us

$$\frac{1}{\pi_1(n)}\left(D_2 - D_1^{(1)}\right) Q_2 = \Pi_2 + A. \tag{1.21}$$

To determine $D_2 - D_1^{(1)}$ we have to derive a matrix $X \in \mathbb{R}^{(n-1)\times(n-1)}$ which solves

$$XQ_2 = \Pi_2 + A. \tag{1.22}$$

Due to the rang-deficiency of Q_2 such a solution X is not unique. Multiplying (iii) in Lemma 1.1.4 with $-A$ yields

$$-AD_2Q_2 = -A\Pi_2 + A$$

where according to Lemma 1.1.2 $-A\Pi_2 = \Pi_2$. Hence, we get $X = -AD_2$ as one possible solution. Because of $\Pi_2 Q_2 = 0$

$$X = -AD_2 + B\Pi_2 \tag{1.23}$$

with B any matrix in $\mathbb{R}^{(n-1)\times(n-1)}$, solves (1.22) as well. Now we get from (1.21), (1.22) and (1.23)

$$\begin{aligned} \frac{1}{\pi_1(n)}\left(D_2 - D_1^{(1)}\right) &= -AD_2 + B\Pi_2 \\ \Leftrightarrow \qquad D_1^{(1)} &= D_2 + \pi_1(n)(AD_2 - B\Pi_2) \end{aligned}$$

and elementwise written

$$d_1(x,y) = d_2(x,y) + \pi_1(n)\left(\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}d_2(z,y) - \sum_{z=1}^{n-1} b(x,z)\pi_2(y)\right), \tag{1.24}$$
$$x,y \in S_2.$$

By inserting (1.24) into (1.17) we obtain

$$\begin{aligned} d_1(x,n) &= \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_2(x,z)\frac{q_1(z,n)}{q_1(n,n)} \\ &\quad -\pi_1(n)\sum_{z=1}^{n-1}\left(\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z) - \sum_{v=1}^{n-1} b(x,v)\pi_2(z)\right)\frac{q_1(z,n)}{q_1(n,n)} \\ &= \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_2(x,z)\frac{q_1(z,n)}{q_1(n,n)} \\ &\quad -\pi_1(n)\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)} \\ &\quad +\frac{\pi_1(n)\sum_{v=1}^{n-1} b(x,v)}{(1-\pi_1(n))q_1(n,n)}\sum_{z=1}^{n-1}\pi_1(z)q_1(z,n), \qquad x \in S_2, \end{aligned}$$

and because $\sum_{z=1}^{n-1} \pi_1(z) q_1(z,n) = -\pi_1(n) q_1(n,n)$ it holds

$$\begin{aligned} d_1(x,n) &= \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_2(x,z) \frac{q_1(z,n)}{q_1(n,n)} \\ &\quad - \pi_1(n) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)} \\ &\quad - \frac{\pi_1(n)^2 \sum_{v=1}^{n-1} b(x,v)}{1-\pi_1(n)} \qquad x \in S_2. \end{aligned} \tag{1.25}$$

From Lemma 1.1.4 (iii) we have

$$\sum_{z=1}^{n} q_1(x,z) d_1(z,y) = \pi_1(y) - \delta(x,y), \qquad x,y \in S_1,$$

which we solve for $d_1(n,y)$ with $x = n$ and $y < n$

$$d_1(n,y) = \frac{\pi_1(y)}{q_1(n,n)} - \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} d_1(z,y). \tag{1.26}$$

Inserting (1.24) into (1.26) gives us

$$\begin{aligned} d_1(n,y) &= \frac{\pi_1(y)}{q_1(n,n)} - \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} d_2(z,y) \\ &\quad - \pi_1(n) \Bigg(\sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,y) \\ &\qquad - \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} \sum_{v=1}^{n-1} b(z,v) \pi_2(y) \Bigg), \qquad y \in S_2. \end{aligned}$$

Since $\sum_{z=1}^{n-1} q_1(n,z) = -q_1(n,n)$ we get

$$\begin{aligned}
d_1(n,y) &= \frac{\pi_1(y)}{q_1(n,n)} - \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} d_2(z,y) + \pi_1(n) \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,y) \\
&\quad + \pi_1(n) \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} \sum_{v=1}^{n-1} b(z,v)\pi_2(y) \\
&= \frac{\pi_1(y)}{q_1(n,n)} - (1-\pi_1(n)) \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} d_2(z,y) \\
&\quad + \pi_1(n) \sum_{z=1}^{n-1} \frac{q_1(n,z)}{q_1(n,n)} \sum_{v=1}^{n-1} b(z,v)\pi_2(y), \qquad y \in S_2. \qquad (1.27)
\end{aligned}$$

Representing (1.15) for $x = y = n$ yields

$$\sum_{z=1}^{n} d_1(n,z) q_1(z,n) = \pi_1(n) - 1$$

which can be solved for $d_1(n,n)$

$$d_1(n,n) = \frac{\pi_1(n) - 1}{q_1(n,n)} - \sum_{z=1}^{n-1} d_1(n,z) \frac{q_1(z,n)}{q_1(n,n)}. \qquad (1.28)$$

By inserting (1.27) into (1.28) we obtain

$$\begin{aligned}
d_1(n,n) &= \frac{\pi_1(n) - 1}{q_1(n,n)} - \sum_{z=1}^{n-1} \frac{\pi_1(z)}{q_1(n,n)} \frac{q_1(z,n)}{q_1(n,n)} \\
&\quad + (1-\pi_1(n)) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)} \\
&\quad - \pi_1(n) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} \sum_{w=1}^{n-1} b(v,w)\pi_2(z) \frac{q_1(z,n)}{q_1(n,n)} \\
&= \frac{2\pi_1(n) - 1}{q_1(n,n)} + (1-\pi_1(n)) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)} \\
&\quad + \frac{\pi_1(n)^2}{1-\pi_1(n)} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} \sum_{w=1}^{n-1} b(v,w). \qquad (1.29)
\end{aligned}$$

Now it remains to compute $\sum_{z=1}^{n-1} b(x,z)$ for $x \in S_2$ what we do by applying (1.24), (1.25) and $-\sum_{z=1}^{n-1} d_1(x,z) = d_1(x,n)$. It holds for $x \in S$

$$
\begin{aligned}
& -\sum_{z=1}^{n-1} d_2(x,z) - \pi_1(n) \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} \sum_{z=1}^{n-1} d_2(v,z) + \pi_1(n) \sum_{z=1}^{n-1} b(x,z) \sum_{v=1}^{n-1} \pi_2(v) \\
&= \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_2(x,z) \frac{q_1(z,n)}{q_1(n,n)} - \pi_1(n) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)} \\
&\quad - \frac{\pi_1(n)^2 \sum_{z=1}^{n-1} b(x,z)}{1-\pi_1(n)}
\end{aligned}
$$

which we solve for $\sum_{z=1}^{n-1} b(x,z)$

$$
\begin{aligned}
\left(\pi_1(n) + \frac{\pi_1(n)^2}{1-\pi_1(n)}\right) \sum_{z=1}^{n-1} b(x,z) &= \frac{\pi_1(n)}{q_1(n,n)} - \sum_{z=1}^{n-1} d_2(x,z) \frac{q_1(z,n)}{q_1(n,n)} \\
&\quad -\pi_1(n) \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} \frac{q_1(n,v)}{q_1(n,n)} d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)}
\end{aligned}
$$

to get

$$
\begin{aligned}
\sum_{z=1}^{n-1} b(x,z) &= \frac{1-\pi_1(n)}{q_1(n,n)} \left(1 - \sum_{z=1}^{n-1} d_2(x,z) \frac{q_1(z,n)}{\pi_1(n)} \right. \\
&\quad \left. - \sum_{z=1}^{n-1} \sum_{v=1}^{n-1} q_1(n,v) d_2(v,z) \frac{q_1(z,n)}{q_1(n,n)} \right), \qquad x \in S_2. \qquad (1.30)
\end{aligned}
$$

We get by inserting (1.30) into (1.24)

$$
\begin{aligned}
d_1(x,y) &= d_2(x,y)+\pi_1(n)\Bigg(\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}d_2(z,y)\\
&\qquad -\frac{1-\pi_1(n)}{q_1(n,n)}\Bigg(1-\sum_{z=1}^{n-1}d_2(x,z)\frac{q_1(z,n)}{\pi_1(n)}\\
&\qquad\qquad -\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}q_1(n,v)d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)}\Bigg)\pi_2(y)\Bigg)\\
&= d_2(x,y)+\pi_1(n)\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}d_2(z,y)-\frac{\pi_1(n)\pi_1(y)}{q_1(n,n)}\\
&\quad +\pi_1(y)\sum_{z=1}^{n-1}d_2(x,z)\frac{q_1(z,n)}{q_1(n,n)}\\
&\quad +\pi_1(n)\pi_1(y)\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)},\quad x,y\in S_2, \qquad (1.31)
\end{aligned}
$$

and into (1.25)

$$
\begin{aligned}
d_1(x,n) &= \frac{\pi_1(n)}{q_1(n,n)}-\sum_{z=1}^{n-1}d_2(x,z)\frac{q_1(z,n)}{q_1(n,n)}\\
&\quad -\pi_1(n)\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)}\\
&\quad -\frac{\pi_1(n)^2}{1-\pi_1(n)}\Bigg(\frac{1-\pi_1(n)}{q_1(n,n)}\Bigg(1-\sum_{z=1}^{n-1}d_2(x,z)\frac{q_1(z,n)}{\pi_1(n)}\\
&\qquad\qquad -\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}q_1(n,v)d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)}\Bigg)\Bigg)\\
&= \frac{\pi_1(n)(1-\pi_1(n))}{q_1(n,n)}-(1-\pi_1(n))\sum_{z=1}^{n-1}d_2(x,z)\frac{q_1(z,n)}{q_1(n,n)} \qquad (1.32)\\
&\quad -\pi_1(n)(1-\pi_1(n))\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)},\qquad x\in S_2,
\end{aligned}
$$

and into (1.27)

$$\begin{aligned}
d_1(n,y) &= \frac{\pi_1(y)}{q_1(n,n)} - (1-\pi_1(n))\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}d_2(z,y) \\
&\quad +\pi_1(n)\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}\Bigg(\frac{1-\pi_1(n)}{q_1(n,n)}\Bigg(1-\sum_{v=1}^{n-1}d_2(z,v)\frac{q_1(v,n)}{\pi_1(n)} \\
&\quad -\sum_{v=1}^{n-1}\sum_{w=1}^{n-1}q_1(n,w)d_2(w,v)\frac{q_1(v,n)}{q_1(n,n)}\Bigg)\Bigg)\pi_2(y) \\
&= \frac{\pi_1(y)(1-\pi_1(n))}{q_1(n,n)} - (1-\pi_1(n))\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}d_2(z,y) \qquad (1.33)\\
&\quad -\pi_1(y)(1-\pi_1(n))\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}\sum_{v=1}^{n-1}d_2(z,v)\frac{q_1(v,n)}{q_1(n,n)}, \qquad y\in S_2,
\end{aligned}$$

and into (1.29)

$$\begin{aligned}
d_1(n,n) &= \frac{2\pi_1(n)-1}{q_1(n,n)} + (1-\pi_1(n))\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)} \\
&\quad +\frac{\pi_1(n)^2}{1-\pi_1(n)}\sum_{z=1}^{n-1}\frac{q_1(n,z)}{q_1(n,n)}\Bigg(\frac{1-\pi_1(n)}{q_1(n,n)}\Bigg(1-\sum_{v=1}^{n-1}d_2(z,v)\frac{q_1(v,n)}{\pi_1(n)} \\
&\quad -\sum_{v=1}^{n-1}\sum_{w=1}^{n-1}q_1(n,w)d_2(w,v)\frac{q_1(v,n)}{q_1(n,n)}\Bigg)\Bigg)
\end{aligned}$$

which simplifies to

$$d_1(n,n) = -\frac{(1-\pi_1(n))^2}{q_1(n,n)} + (1-\pi_1(n))^2\sum_{z=1}^{n-1}\sum_{v=1}^{n-1}\frac{q_1(n,v)}{q_1(n,n)}d_2(v,z)\frac{q_1(z,n)}{q_1(n,n)}. \qquad (1.34)$$

That this solution D_1 is indeed the uniquely determined deviation matrix follows by the fact that it simultaneously solves $\Pi_1 D_1 = 0$ and $Q_1 D_1 = \Pi_1 - I$; see Lemma 1.4.2.

Now that we have finally derived a recursive formula of the deviation matrix, we present the corresponding algorithm. First, we have to compute

the transition rates for every state space S_k, $k \in \{2, \dots, n\}$, starting with a given generator matrix Q_1 we successively compute the transition rates of the smaller state spaces applying (1.12). In the literature (see, e.g., [126]), this initial step is known as *reduction*. Afterwards the *recovery* step begins, coming from $D_n = (0)$ and $\Pi_n = (1)$, with the successive computation of the limiting matrices by applying the recursive formula provided in Remark 1.4.3. Finally, Equations (1.31) to (1.34) yield the deviation matrix.
We will illustrate this algorithm by applying it to the M/M/1/N queue.

Example 1.4.4 (Deviation Matrix of an M/M/1/N Queue) *Let $\mathcal{X}_1$ be the continuous-time Markov process describing the number of customers waiting in a single-server queue having a maximal waiting room capacity of* 2 *so that $S_1 = \{0, 1, 2\}$. Customers arrive at the queue according to a Poisson(λ) process and are served for an exponentially-μ distributed service time. Such a process has generator matrix*

$$Q_1 = \begin{pmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda+\mu) & \lambda \\ 0 & \mu & -\mu \end{pmatrix}$$

and by (1.12) we get

$$Q_2 = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \qquad \text{and} \qquad Q_3 = (0)$$

with $S_2 = \{0, 1\}$ and $S_3 = \{0\}$. Coming from

$$\Pi_3 = (1) \qquad \text{and} \qquad D_3 = (0)$$

we apply Remark 1.4.3 to get

$$\Pi_2 = \begin{pmatrix} \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \\ \frac{\mu}{\lambda+\mu} & \frac{\lambda}{\lambda+\mu} \end{pmatrix}$$

and

$$\Pi_1 = \begin{pmatrix} \frac{\lambda\mu^2+\mu^3}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2\mu+\lambda\mu^2}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2}{(\lambda+\mu)^2-\lambda\mu} \\ \frac{\lambda\mu^2+\mu^3}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2\mu+\lambda\mu^2}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2}{(\lambda+\mu)^2-\lambda\mu} \\ \frac{\lambda\mu^2+\mu^3}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2\mu+\lambda\mu^2}{(\lambda+\mu)^3-\lambda\mu(\lambda+\mu)} & \frac{\lambda^2}{(\lambda+\mu)^2-\lambda\mu} \end{pmatrix}$$

and with (1.31), (1.32), (1.33) and (1.34) we successively compute the deviation matrix

$$D_2 = \begin{pmatrix} \frac{\lambda}{(\lambda+\mu)^2} & -\frac{\lambda}{(\lambda+\mu)^2} \\ -\frac{\mu}{(\lambda+\mu)^2} & \frac{\mu}{(\lambda+\mu)^2} \end{pmatrix}$$

and

$$D_1 = \begin{pmatrix} \frac{\lambda\left(\mu^2+3\lambda\mu+\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} & -\frac{\lambda\left(\mu^2+\lambda\mu-\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} & -\frac{2\lambda^2(\lambda+\mu)}{(\lambda^2+\lambda\mu+\mu^2)^2} \\ -\frac{\mu\left(\mu^2+\lambda\mu-\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} & \frac{\lambda^3+\mu^3}{(\lambda^2+\lambda\mu+\mu^2)^2} & \frac{\lambda\left(\mu^2-\lambda\mu-\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} \\ -\frac{2\mu^2(\lambda+\mu)}{(\lambda^2+\lambda\mu+\mu^2)^2} & \frac{\mu\left(\mu^2-\lambda\mu-\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} & \frac{\mu\left(\mu^2+3\lambda\mu+\lambda^2\right)}{(\lambda^2+\lambda\mu+\mu^2)^2} \end{pmatrix}.$$

Note that this result is in accordance with Equation (1.8).

1.4.3 Mean Passage Times

An alternative formula which provides a closed-form representation of the individual entries of the deviation matrix was presented in [32] for a uni-chain continuous-time Markov process $\mathcal{X}$ with the derivation provided in Chapter 2 of this thesis. The formula

$$d(x, y) = \pi(y) \left(m_\pi(y) - m(x, y)\right), \qquad x, y \in S, \tag{1.35}$$

is especially useful when state space S is large or infinite and according to [135] even applies for continuous state spaces S (see also Section 1.6 of this thesis). However, it requires the knowledge of the mean passage times which are indeed not necessarily given or easy to derive. Therefore, (1.35) is restricted to processes $\mathcal{X}$ for which $m(x, y)$ is computable in closed-form. For examples of the applicability of (1.35) we refer to Section 1.6 as well as Chapter 2, 4 and 5 where we use it to compute the deviation matrix of various Markov processes.

Example 1.4.5 (Deviation Matrix of an M/M/c+M Queue) *Let $\mathcal{X}$ on $S \stackrel{\text{def}}{=} \mathbb{N}_0$ describe the number of customers waiting in a queue which they entered according to a Poisson(λ) process. Inside this system, customers are either served at one of c servers each for an exponentially-μ distributed service time or they are queueing until their service starts or they become impatient and abandon. Their patience expires after an exponentially-α distributed period. An overview of this system introduced by Palm in [109], also known as Erlang-A queue or in short notation M/M/c+M, is provided in Figure 1.2. Such a system has a Q-matrix*

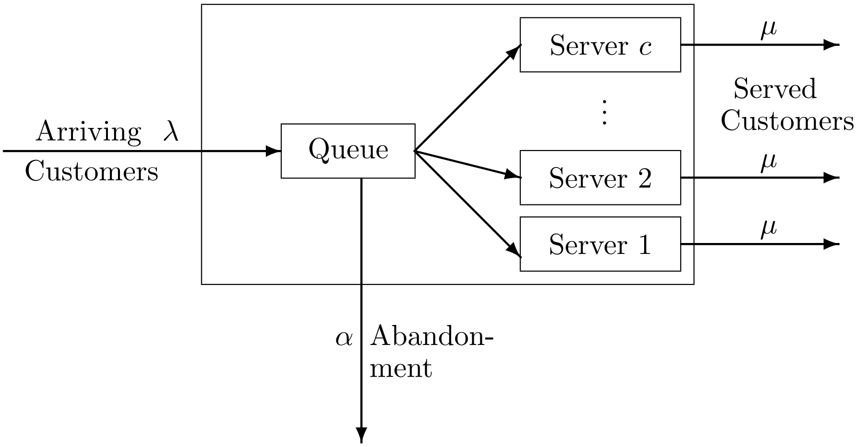

Figure 1.2: Structure of an M/M/c+M Queueing System with Abandonment

$$q(x,y) = \begin{cases} \lambda & y = x+1, \\ \min\{c,x\}\mu + \max\{0, x-c\}\alpha & y = x-1, \\ -\lambda - \min\{c,x\}\mu - \max\{0,x-c\}\alpha & y = x, \\ 0 & \text{otherwise,} \end{cases}$$

and the stationary distribution is given by

$$\pi(y) = K^{-1} \frac{\lambda^{\min\{y,c\}}}{\min\{y,c\}!\mu^{\min\{y,c\}}} \prod_{x=1}^{y-c} \frac{\lambda}{c\mu + x\alpha}, \qquad y \in S, \tag{1.36}$$

where an empty product is supposed to be 1 *and the normalizing constant is given by*

$$K = \sum_{x=0}^{c} \frac{\lambda^x}{x!\mu^x} + \frac{\lambda^c}{c!\mu^c} \sum_{y=c+1}^{\infty} \prod_{x=1}^{y-c} \frac{\lambda}{c\mu + x\alpha}.$$

We will now exemplary compute the deviation matrix entries $d(x,y)$ *for* $y = 0$*. Using the representation of mean first passage times provided in [80], p.248, we get for* $x \leq c$

$$m(x,0) = \sum_{z=0}^{x-1} \frac{z!\mu^z}{\lambda^{z+1}} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right) \tag{1.37}$$

and for $x > c$

$$\begin{aligned} m(x,0) &= \sum_{z=0}^{c-1} \frac{z!\mu^z}{\lambda^{z+1}} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right) \qquad (1.38) \\ &\quad + \sum_{z=c+1}^{x-1} \left(\frac{1}{\lambda} \prod_{v=1}^{z-c} \frac{c\mu + v\alpha}{\lambda} \right) \left(\sum_{w=z+1}^{\infty} \prod_{v=1}^{w-c} \frac{\lambda}{c\mu + v\alpha} \right) \end{aligned}$$

and

$$\begin{aligned} m_\pi(0) &= \frac{1}{\lambda K} \left(\sum_{z=0}^{c} \frac{z!\mu^z}{\lambda^z} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right)^2 \right. \qquad (1.39) \\ &\quad \left. + \frac{\lambda^c}{c!\mu^c} \sum_{z=c+1}^{\infty} \left(\prod_{v=1}^{z-c} \frac{c\mu + v\alpha}{\lambda} \right) \left(\sum_{w=z+1}^{\infty} \prod_{v=1}^{w-c} \frac{\lambda}{c\mu + v\alpha} \right)^2 \right) \end{aligned}$$

which is finite whenever $\lambda < \infty$ so that the existence of the deviation matrix is assured for bounded arrival rates. Inserting (1.36), (1.37), (1.38) and (1.39) into (1.35) yields the first column of the deviation matrix for $x \leq c$

$$\begin{aligned} d(x,0) &= \frac{1}{\lambda K} \left(\frac{1}{K} \sum_{z=0}^{c} \frac{z!\mu^z}{\lambda^z} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right)^2 - \sum_{z=0}^{x-1} \frac{z!\mu^z}{\lambda^z} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right) \right. \\ &\quad \left. + \frac{\lambda^c}{Kc!\mu^c} \sum_{z=c+1}^{\infty} \left(\prod_{v=1}^{z-c} \frac{c\mu + v\alpha}{\lambda} \right) \left(\sum_{w=z+1}^{\infty} \prod_{v=1}^{w-c} \frac{\lambda}{c\mu + v\alpha} \right)^2 \right) \end{aligned}$$

and for $x > c$

$$\begin{aligned} d(x,0) &= \frac{1}{\lambda K} \left(\frac{1}{K} \sum_{z=0}^{c} \frac{z!\mu^z}{\lambda^z} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right)^2 - \sum_{z=0}^{c-1} \frac{z!\mu^z}{\lambda^z} \left(K - \sum_{v=0}^{z} \frac{\lambda^v}{v!\mu^v} \right) \right. \\ &\quad + \frac{\lambda^c}{Kc!\mu^c} \sum_{z=c+1}^{\infty} \left(\prod_{v=1}^{z-c} \frac{c\mu + v\alpha}{\lambda} \right) \left(\sum_{w=z+1}^{\infty} \prod_{v=1}^{w-c} \frac{\lambda}{c\mu + v\alpha} \right)^2 \\ &\quad \left. - \sum_{z=c+1}^{x-1} \left(\prod_{v=1}^{z-c} \frac{c\mu + v\alpha}{\lambda} \right) \left(\sum_{w=z+1}^{\infty} \prod_{v=1}^{w-c} \frac{\lambda}{c\mu + v\alpha} \right) \right). \end{aligned}$$

1.4.4 Taboo Representation

Let $\tilde{\mathcal{X}}$ be a discrete-time Markov chain on state space S which might be of multi-chain structure. Then according to [66] the deviation matrix can be expressed by

$$\tilde{D} = (I - \tilde{\Pi}) \sum_{n=0}^{\infty} {}_B\tilde{P}^n (I - \tilde{\Pi}) \tag{1.40}$$

where $B \subset S$ is a (finite) taboo set of states consisting of representatives from each positive recurrent class and

$$\left({}_B\tilde{P}\right)_{(x,y)} = \begin{cases} p(x,y) & x \in S,\ y \notin B, \\ 0 & x \in S,\ y \in B. \end{cases}$$

The existence of $\tilde{D}$ is assured by the following condition

$$\left\| {}_B\tilde{P} \right\|_{v} < 1$$

for some v-norm. In case of finite state space S we can furthermore transform (1.40) to

$$\tilde{D} = (I - \tilde{\Pi}) \left(I - {}_B\tilde{P}\right)^{-1} (I - \tilde{\Pi}). \tag{1.41}$$

Let $\mathcal{X}$ be a uniformizable continuous-time Markov process on S having generator matrix Q. With (1.7) we can associate to any infinitesimal generator Q having ν-bounded $q(x)$, $x \in S$, a transition matrix

$$\tilde{P} = I + \frac{1}{\nu} Q$$

and therefore apply (1.40) as well. We will illustrate (1.40) now by applying it to a simple example.

Example 1.4.6 *Let* $S = \{1, 2\}$ *and*

$$\tilde{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

with $p, q \in (0,1)$ *then it holds for the* n*-step transition probability matrix*

$$\tilde{P}(n) = \frac{1}{p+q} \left(\begin{pmatrix} q & p \\ q & p \end{pmatrix} + (1-p-q)^n \begin{pmatrix} p & -p \\ -q & q \end{pmatrix} \right).$$

Since $|1-p-q|<1$ *we get for the limiting distribution*

$$\tilde{\Pi} = \lim_{n\to\infty} \tilde{P}(n) = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix}.$$

The deviation matrix can be computed using (1.5) what gives us

$$\begin{aligned}\tilde{D} &= \sum_{n=0}^{\infty} \frac{(1-p-q)^n}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix} = \frac{1}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\sum_{n=0}^{\infty}(1-p-q)^n \\ &= \frac{1}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\frac{1}{1-(1-p-q)} = \frac{1}{(p+q)^2}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.\end{aligned}$$

Now suppose that state 2 *becomes a taboo state so that*

$${}_2\tilde{P} = \begin{pmatrix} 1-p & 0 \\ q & 0 \end{pmatrix}$$

and by (1.41) we get

$$\begin{aligned}\tilde{D} &= \frac{1}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\begin{pmatrix} p & 0 \\ -q & 1 \end{pmatrix}^{-1}\frac{1}{p+q}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix} \\ &= \frac{1}{(p+q)^2}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\end{aligned}$$

which is obviously equal to the result we obtained by the direct computation using (1.5).

1.5 Interpretation

At first glance, the deviation matrix measures the speed with which the transition matrix $P(t)$ approaches the limiting distribution Π. Due to the following property of the limiting distribution

$$\sum_{z\in S}\pi(x,z)p(z,y;t) = \pi(x,y), \qquad x,y\in S,$$

we can rewrite the definition of the deviation matrix given in (1.3) so that

$$\begin{aligned}d(x,y) &= \int_0^\infty p(x,y;t) - \pi(x,y)dt \\ &= \int_0^\infty p(x,y;t) - \sum_{z\in S}\pi(x,z)p(z,y;t)dt, \qquad x,y\in S.\end{aligned}$$

Now it is easy to see that the deviation matrix measures the difference between starting in state x of some ergodic class $S_1 \subseteq S$ and going to y, compared to starting initially stationary within class S_1 and going to y. However, whenever the transition probabilities $p(x, y; t)$, $x, y \in S$, are not monotone in t so that $p(x, y; t) - \pi(x, y)$ might be neither strictly positive nor negative the entries partially cancel out and the value $d(x, y)$ might overestimate the speed of convergence. We illustrate this drawback in the following simple discrete-time example.

Example 1.5.1 *Recall Example 1.4.6 where* $S = \{1, 2\}$ *and*

$$\tilde{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

and

$$\tilde{P}(n) = \frac{1}{p+q}\left(\begin{pmatrix} q & p \\ q & p \end{pmatrix} + (1-p-q)^n \begin{pmatrix} p & -p \\ -q & q \end{pmatrix}\right)$$

where $p, q > 0$ *assures that* $\tilde{\mathcal{X}}$ *is a uni-chain and it holds*

$$\tilde{\Pi} = \frac{1}{p+q}\begin{pmatrix} q & p \\ q & p \end{pmatrix} \qquad \text{and} \qquad \tilde{D} = \frac{1}{(p+q)^2}\begin{pmatrix} p & -p \\ -q & q \end{pmatrix}.$$

Setting $p_1 = q_1 = 0.5$ *yields the limiting distribution* $\pi_1(1) = \pi_1(2) = 0.5$ *and the deviation matrix*

$$D_1 = \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}.$$

While $p_2 = q_2 = 0.9$ *lead to the same limiting distribution* $\pi_2(1) = \pi_2(2) = 0.5$ *the deviation matrix is given by*

$$D_2 = \begin{pmatrix} 0.28 & -0.28 \\ -0.28 & 0.28 \end{pmatrix}.$$

Although the two cases have the same limiting distributions what stems from the fact that $p = q$ *in both cases, the entries of their deviation matrices diverge significantly. Although the second deviation matrix has smaller entries* 0.28 *the difference in between* $\tilde{p}_2(x, y; n) - \tilde{\pi}_2(y)$ *is larger than zero for any finite* n *whereas* $\tilde{p}_1(x, y; n) - \tilde{\pi}_1(y)$ *is equal to zero for* $n > 0$ *as* $(1 - p_1 - q_1)^n = 0^n$. *Therefore,* $\tilde{p}_1(x, y; n)$ *reaches its limiting value* $\tilde{\pi}_1(y)$ *within one step whereas* $\tilde{p}_2(x, y; n)$ *approaches* $\tilde{\pi}_2(y)$ *more slowly. A fact which is overlooked by the*

deviation matrix since $\tilde{p}_2(x, y; n) - \tilde{\pi}_2(y)$ *is alternating due to* $(1 - p_2 - q_2)^n = (-0.8)^n$ *so that positive and negative values eliminate each other.*
However, if we transform $\tilde{\mathcal{X}}$ *to the continuous-time process* $\mathcal{X}$ *with*

$$Q = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix}$$

from what we get

$$P(t) = exp[Qt] = \frac{1}{p+q} \begin{pmatrix} q + pe^{-(p+q)t} & p - pe^{-(p+q)t} \\ q - qe^{-(p+q)t} & p + qe^{-(p+q)t} \end{pmatrix}$$

and the limiting matrix

$$\Pi = \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

and deviation matrix computed using (1.10)

$$\begin{aligned} D &= \left(\frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} - \begin{pmatrix} -p & p \\ q & -q \end{pmatrix} \right)^{-1} - \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} \\ &= \frac{1}{(p+q)^2} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix} \end{aligned}$$

where $\tilde{\Pi} = \Pi$ *and* $\tilde{D} = D$ *is in accordance with our findings of Section 1.2, we can easily see that* $P(t)$ *approaches* Π *monotonously and* $p(x, y; t) - \pi(y)$ *is not alternating at all.*

To avoid wrong conclusions concerning the speed of convergence whenever monotonicity of $p(x, y; t)$ in t cannot be assured in advance, we recommend to replace $d(x, y)$ by $\int_0^\infty |p(x, y; t) - \pi(x, y)| dt$, i.e., integrate over the absolute values of $p(x, y; t) - \pi(x, y)$ instead.
Beside its application as a convergence measure the property of the deviation matrix of comparing the transition from x to y with starting stationary and going to y is useful to many applications. We will stick now to the unichain case from which the more general case of multi-chains can be directly derived. Multiplying D with a cost vector c gives us (provided the interchange of summation and integration is allowed)

$$\sum_{y \in S} d(x, y) c(y) = \int_0^\infty \sum_{y \in S} p(x, y; t) c(y) - \sum_{y \in S} \pi(y) c(y) dt, \qquad x \in S,$$

where $\sum_{y \in S} \pi(y)c(y)$ are the average long-term costs of $\mathcal{X}$ whereas $\int_0^\infty \sum_{y \in S} p(x,y;t)c(y)dt$ are the costs which occur when initially starting in x. Whenever $\sum_{y \in S} d(x,y)c(y) < 0$ starting in x is superior to the current system's long-term behavior. The theory of Markov decision processes (MDPs) capitalizes on this fact calling Dc the *value function* or *bias vector*. (for a detailed introduction to MDPs see, e.g., Chapter 2, 4 and 5 and the references therein).

While this thesis is devoted to the exact computation of D and an analysis of the deviation matrix based on its properties as a matrix, an alternative approach is mainly pursued by Cao et al.(see, e.g., [23, 25, 26, 27]). Focussing on a finite state space discrete-time Markov chains, their view is simulation based for they regard individual sample paths $\{\tilde{X}_0, \tilde{X}_1, \dots\}$ instead of investigating $\tilde{D}$ as a given matrix. Coming from

$$\tilde{D}c(x) = \lim_{N \to \infty} E\left[\sum_{n=0}^{N-1} \left(c(\tilde{X}_n) - \tilde{\pi}c\right) \middle| \tilde{X}_0 = x\right], \qquad x \in S, \tag{1.42}$$

(note that there is a slight variation in notation between the cited papers and the one used herein, compare, e.g., (2) in [25]) they get the deviation matrix entry $\tilde{d}(x,y)$ by setting $c(z) \stackrel{\text{def}}{=} \mathbf{1}_{\{z=y\}}$ and in case that c represents the previously introduced cost vector they arrive at the value function which they call *performance potential* due to its property of measuring the potential contribution of state x to the long-run average performance πc.
Cao's approach is much more intuitively. He uses the fact that with the strong Markov property it holds that two irreducible processes $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{X}}'$ on finite state space S having the same transition matrix $\tilde{P}$ but being started in x and x', respectively, will couple after random time and afterwards behave statistically identical. Therefore, any ergodic process $\tilde{\mathcal{X}}$ will approach its limiting matrix $\tilde{\Pi}$ after a random number of steps, say L, so that with (1.42) it follows

$$E\left[\sum_{n=L}^{\infty} \left(c(\tilde{X}_n) - \tilde{\pi}c\right) \middle| \tilde{X}_0 = x\right] = 0, \qquad x \in S.$$

Furthermore,

$$E\left[\sum_{n=0}^{L-1} \left(c(\tilde{X}_n) - \tilde{\pi}c\right) \middle| \tilde{X}_0 = x\right] = \tilde{D}c(x), \qquad x \in S, \tag{1.43}$$

(1.43) measures the cost (dis)advantage that arises until process $\tilde{\mathcal{X}}$ started in x couples with its stationary version.
In their recent paper [28] Cao and Zhang proposed the k^{th} bias vector, defined by

$$\tilde{D}^{(k)}c(x) \stackrel{\text{def}}{=} (-1)^k \sum_{n=0}^{\infty} \binom{k+n}{k} E\left[\left(c(\tilde{X}_n) - \tilde{\pi}c\right)\Big| \tilde{X}_0 = x\right], \qquad x \in S,$$

and transfered to matrix notation

$$\begin{aligned}\tilde{D}^{(k)}c &\stackrel{\text{def}}{=} (-1)^k \sum_{n=0}^{\infty} \binom{k+n}{k} \left(\tilde{P}^n - \tilde{\Pi}\right) c \\ &= -\left(I - \tilde{P} + \tilde{\Pi}\right)^{-1} \tilde{D}^{(k-1)}c, \qquad k \geq 1, \end{aligned} \tag{1.44}$$

where $\tilde{D}^{(0)} = \tilde{D}$, as a tool to iteratively derive $(k-1)$-discount optimal policies f (for an introduction to discount optimal policies we refer to [45, 113]). In [140] they transferred their findings to a continuous-time setting with

$$D^{(k)}c \stackrel{\text{def}}{=} -\int_0^{\infty} \left(P(t) - \Pi\right) D^{(k-1)}c dt = (-1)^k D^{k+1}c, \qquad k \geq 1,$$

where $D^{(0)} = D$. Complementary, to the set of average performance optimal policies F_0 and the set of bias optimal policies F_1 chosen from a given set of policies F which are investigated in this thesis (see Chapter 2, 4 and 5) and satisfy

$$\operatorname{argmin}_{f \in F} \{\pi_f c\} \qquad \text{and} \qquad \operatorname{argmin}_{f \in F_0} \{D_f c\},$$

respectively, they state in [28, 140] that for any k^{th} bias optimal policy f, i.e., f satisfying

$$\operatorname{argmin}_{f \in F_{k-1}} \left\{D_f^{(k)} c\right\} \tag{1.45}$$

where $F_{k-1} \subseteq F$ denotes the set of policies being $(k-1)$-bias optimal, it holds that

$$f = \operatorname{argmin}_f \left\{\sum_{n=0}^{k} \lambda^{n-1} D_f^{(n)} c\right\} \tag{1.46}$$

where $\lambda > 0$ is some discount factor, so that f is $(k-1)$-discount optimal as well. Furthermore, for $k \to \infty$ the sum in (1.46) becomes a Laurent series and with Equation (8) in [140] we get

$$f = \text{argmin}_f \left\{ \sum_{n=0}^{\infty} \lambda^{n-1} D_f^{(n)} c \right\} = \text{argmin}_f \left\{ \int_0^{\infty} e^{-\lambda t} P_f(t) c \, dt \right\}$$

so that any ∞-bias optimal policy is yet discount optimal. Additionally, from the definition of k^{th} bias optimality in (1.45) we know that any k^{th} bias optimal policy is $(k-1)$-bias optimal as well. But while simple bias optimality assesses future costs equally, higher order optimality puts more weight on short-term costs whereas long-run expenses have less influence on the decision, an approach which is in accordance with the conventional application of discount factors in decision making.

1.6 Extension to Continuous State Spaces

Continuous-time Markov processes on general state spaces, especially on the line of real numbers, are a common tool in operations research and physics. Their ergodicity properties were extensively investigated by Borovkov, Hernández-Lerma and Lasserre as well as Meyn and Tweedie (see, e.g., [20, 61, 96, 94, 95]) whereas optimization via the theory of Markov decision processes (MDPs) in discrete-time as well as in continuous-time was presented in [93, 118].

However, to efficiently optimize an MDP $\mathcal{X}_f$ over its set of possible policies f requires the knowledge of the so-called value function or bias vector; see, e.g., [113]. As the value function is given by Dc, where c denotes a cost vector and D the deviation operator, it is especially useful to know D explicitly. Exemplary, we will derive the closed-form representation of D for a Brownian motion.

First of all, we have to broaden the concept of Section 1.1 to the setting of continuous state spaces:

Let $\mathcal{X} \stackrel{\text{def}}{=} \{X_t : t \in \mathbb{R}_+\}$ be a time-homogeneous Markov process on a continuous state space $S \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, with associated Borel σ-field $\mathcal{S}$. Such a process is determined by its transition probability kernel $P(t)$ with

$$P(x, A; t) \stackrel{\text{def}}{=} \mathcal{P}\{X_t \in A | X_0 = x\}, \qquad x \in S, \ A \in \mathcal{S},$$

which acts on measurable functions $f : S \to \mathbb{R}$ via

$$P(t)f(\cdot) = \int_S P(\cdot, dy; t) f(y), \qquad t \geq 0,$$

such that the integral exists, and on measures via

$$\mu P(t) = \int_S \mu(dy) P(y, \cdot; t), \qquad t \geq 0,$$

(the existence of $\mathcal{X}$ determined by $P(t)$ and a prescribed initial distribution follows by Kolmogorov's extension theorem). By $\mathcal{C}_v(S)$ we denote the set of all real-valued measurable functions f that satisfy

$$\|f\|_v < \infty \tag{1.47}$$

where $\|\cdot\|_v$ is the weighted supremum norm, also called v-norm, defined by

$$\|f\|_v \stackrel{\text{def}}{=} \sup_{x \in S} \frac{|f(x)|}{v(x)}$$

for any $v : \ S \to [1, \infty]$ with $v(x_0) < \infty$ for at least one $x_0 \in S$ chosen appropriately. Then $(\mathcal{C}_v(S), \|\cdot\|_v)$ forms a Banach space; see [48], p.921. Analogously, we define the v-norm of a measure by

$$\|\mu\|_v \stackrel{\text{def}}{=} \sup_{\|g\|_v \leq 1} \int_S |g(y)| \, \mu(dy)$$

and of a kernel by

$$\|P(t)\|_v \stackrel{\text{def}}{=} \sup_{\|w\|_v \leq 1} \frac{\|P(t)w\|_v}{\|w\|_v}.$$

Throughout this section we will assume all processes $\mathcal{X}$ to be *positive Harris recurrent* thus ensuring the existence of a unique invariant probability measure π on $\mathcal{S}$ satisfying

$$\pi(A) = \int_S \pi(dy) P(y, A; t), \qquad \forall A \in \mathcal{S}, \ t \geq 0, \tag{1.48}$$

with the associated kernel Π, a bounded linear operator on $\mathcal{C}_v(S)$, $v \geq 1$, with $\Pi(x, A) = \pi(A)$ with $x \in S$ and $A \in \mathcal{S}$ (see, e.g., [95], p.491, and the references mentioned therein). If we additionally suppose $\mathcal{X}$ to be an aperiodic process (where aperiodicity for continuous-time and continuous state space processes is defined according to [41], p.1675) and to satisfy the geometric drift criterion of Theorem 6.2 in [41], p.1683, (a condition that is fulfilled by all processes appearing in this section) we will get by Theorem 7.4 in [41], p.1689, that

$$\|P(t) - \Pi\|_v \leq c\delta^t, \qquad t \geq 0, \tag{1.49}$$

for some $c < \infty$, $\delta \in (0,1)$ and $v \geq 1$. That is, our process $\mathcal{X}$ is v-uniform ergodic. We define the deviation operator D (also known as the fundamental kernel, see, e.g., p.316f in [82]), a linear operator $D : \mathcal{C}_v(S) \to \mathcal{C}_v(S)$ for $v \geq 1$, with

$$D(x, A) = \int_0^\infty P(x, A; s) - \pi(A) ds, \qquad x \in S, \ A \in \mathcal{S},$$

of which the existence can be directly derived from our somewhat stronger assumption of finite $\|D\|_v$ which follows from (1.49) via

$$\|D\|_v \leq \int_0^\infty \|P(s) - \Pi\|_v ds \leq c \int_0^\infty \delta^s ds = -\frac{c}{\ln[\delta]} < \infty. \tag{1.50}$$

We now suppose that $\mathcal{X}$ is a diffusion process defined on state space $S = [l, r]$ or $[l, r)$ where l is the lower boundary and r the upper boundary of a continuous interval and $\mathcal{X}$ is determined by its drift coefficient $\mu(x)$ and diffusion coefficient $\sigma^2(x)$, $x \in S$, as well as its behavior at the boundary points l, r. We assume that the transition probability kernels of all diffusion processes appearing in this section have Lebesgue densities $p(x, y; t)$, $x, y \in S$, as their infinitesimal coefficients are supposed to be smooth. For a sufficient condition on $\mu(x)$ and $\sigma^2(x)$ we refer to [73], p.213f, as well as Theorem 3 in [92]. It follows that the limiting distribution has Lebesgue density too which is according to (6.22) in [73], p.241, given by

$$\pi(y) = \frac{\eta(y)}{\int_l^r \eta(z) dz}, \qquad y \in S, \tag{1.51}$$

with

$$\eta(x) = \frac{\exp\left[\int_l^x \frac{2\mu(z)}{\sigma^2(z)} dz\right]}{\sigma^2(x)}, \qquad x \in S. \tag{1.52}$$

Note that the usage of π for the Lebesgue density is a slight abuse of notation as in (1.48) we introduced π as the limiting distribution itself. Under the appropriate conditions mentioned above the existence of a Lebesgue density for the limiting distribution is assured so that in what follows we can stick to the density and use the expressions $\pi(y)$, $y \in S$, for the density instead of the respective distribution.

Lemma 1.6.1 *The density of the deviation operator D of a diffusion process is given by*

$$d(x, y) \stackrel{\text{def}}{=} \int_0^\infty p(x, y; s) - \pi(y) ds, \qquad x, y \in S.$$

Proof: From the definition of the deviation operator we get

$$\begin{aligned} D(x,A) &= \int_0^\infty P(x,A;s) - \pi(A)ds = \int_0^\infty \left(\int_A p(x,y;s)dy - \int_A \pi(y)dy \right) ds \\ &= \int_0^\infty \int_A \left(p(x,y;s) - \pi(y) \right) dyds, \qquad x \in S,\ A \in \mathcal{S}. \end{aligned}$$

Justified by Fubini's theorem and (1.49), we can interchange the order of integration and get

$$D(x,A) = \int_A \int_0^\infty \left(p(x,y;s) - \pi(y) \right) dsdy = \int_A d(x,y)dy, \qquad x \in S,\ A \in \mathcal{S}.$$

■

For the computation of $d(x,y)$, $x,y \in S$, we will apply a formula based on mean first passage times which was introduced in [32] and proven to hold for denumerable state spaces (see Section 1.4.3 for the presentation of this formula and Chapter 2 for the respective proof on denumerable state space).

Theorem 1.6.2 *If the deviation operator D of a continuous-time Markov process $\mathcal{X}$ defined on an interval S exists its density is represented by*

$$d(x,y) = \pi(y) \left(\int_S \pi(z) m(z,y) dz - m(x,y) \right), \qquad x,y \in S, \tag{1.53}$$

where $m(x,y)$ denotes the mean first passage time from state x to y.

It was hinted at by Whitt in [135] that this passage time formula for the density of the deviation operator carries over to diffusion processes; see (46) and (47) in [135]. However, he does not provide a proof of this conjecture. Therefore, we will close this gap herein.

Proof: Let

$$T_y = \begin{cases} \inf\{t \geq 0 \,|\, X_t \geq y\} & X_0 \leq y, \\ \inf\{t \geq 0 \,|\, X_t \leq y\} & X_0 > y, \end{cases} \qquad t \geq 0,$$

be the first passage time to $y \in S$ and let

$$F(x,y;t) = \mathcal{P}\{T_y \leq t \,|\, X_0 = x\}, \qquad x,y \in S,$$

denote its associated distribution function for $\mathcal{X}$ started in state $x \in S$. Furthermore, we introduce the short-hand notation

$$F_\pi(y;t) = \int_S \pi(x)F(x,y;t)dx, \qquad y \in S, \tag{1.54}$$

which represents the distribution function of the first passage time to y for $\mathcal{X}$ started initially stationary. According to [110, 119] it holds for the density of the transition probability function of a diffusion process $\mathcal{X}$

$$p(x,y;t) = \int_0^t p(y,y;t-s)F(x,y;ds), \qquad x \neq y, \tag{1.55}$$

and by replacing $p(x,y;t)$ in the righthand side of (1.48) (an equation which was introduced for distributions but holds for densities as well) by (1.55) we get

$$\pi(y) = \int_S \pi(x) \int_0^t p(y,y;t-s)F(x,y;ds)dx, \qquad y \in S,$$

where we can interchange the order of integration justified by Fubini's theorem and we get with (1.54)

$$\pi(y) = \int_0^t p(y,y;t-s)F_\pi(y;ds), \qquad y \in S.$$

Applying the Laplace transformation to this equation (for a broader investigation of Laplace transformations of diffusion processes we refer to p.149ff in [17]) yields

$$\pi(y)\int_0^\infty e^{-\alpha t}dt = \int_0^\infty e^{-\alpha t}\int_0^t p(y,y;t-s)F_\pi(y;ds)dt, \qquad y \in S,\ \alpha > 0, \tag{1.56}$$

which we can transform considering the fact that the Laplace transformation of the occurring convolution equals the product of the Laplace transformation of the function $p(y,y;t)$ and the Laplace-Stieltjes transformation of the measure $F_\pi(y;t)$. By using the notation $\hat{}$ and replacing time parameter t by α to reflect Laplace transformations within the following equations we get for (1.56)

$$\frac{\pi(y)}{\alpha} = \hat{p}(y,y;\alpha)\hat{F}_\pi(y;\alpha), \qquad y \in S,\ \alpha > 0,$$

which we solve for $\hat{p}(y,y;\alpha)$

$$\hat{p}(y,y;\alpha) = \frac{\pi(y)}{\alpha \hat{F}_\pi(y;\alpha)}, \qquad y \in S,\ \alpha > 0. \tag{1.57}$$

Now we introduce

$$\hat{d}(y,y;\alpha) \stackrel{\text{def}}{=} \int_0^\infty e^{-\alpha t}(p(y,y;t) - \pi(y))dt = \hat{p}(y,y;\alpha) - \frac{\pi(y)}{\alpha}, \qquad y \in S,\ \alpha > 0,$$

and with (1.57) it holds

$$\hat{d}(y,y;\alpha) = \frac{\pi(y)(1-\hat{F}_\pi(y;\alpha))}{\alpha \hat{F}_\pi(y;\alpha)}, \qquad y \in S,\ \alpha > 0.$$

As we are especially interested in the limit $\alpha \to 0$ we apply l'Hospital's rule which yields

$$\lim_{\alpha\to 0} \hat{d}(y,y;\alpha) = \lim_{\alpha \to 0} \frac{-\pi(y)\int_0^\infty (-t)e^{-\alpha t}F_\pi(y;dt)}{\int_0^\infty e^{-\alpha t}F_\pi(y;dt) + \alpha \int_0^\infty (-t)e^{-\alpha t}F_\pi(y;dt)}, \qquad y \in S.$$

Due to $\int_0^\infty F_\pi(dt) = 1$ and $\int_0^\infty tF_\pi(dt) = \int_S \pi(z)m(z,y)dz$, and the fact that a Tauberian theorem justifies $\lim_{\alpha\to 0}\hat{d}(y,y;\alpha) = d(y,y)$ we get

$$d(y,y) = \pi(y)\int_S \pi(z)m(z,y)dz, \qquad y \in S.$$

We define

$$\hat{d}(x,y;\alpha) = \int_0^\infty e^{-\alpha t}(p(x,y;t) - \pi(y))dt, \qquad x,y \in S,$$

and directly get

$$\begin{aligned}\hat{d}(x,y;\alpha) &= \int_0^\infty e^{-\alpha t}(p(y,y;t) - \pi(y) + p(x,y;t) - p(y,y;t))dt \\ &= \hat{d}(y,y;\alpha) + \hat{p}(x,y;\alpha) - \hat{p}(y,y;\alpha), \qquad x,y \in S. \end{aligned} \tag{1.58}$$

Applying the Laplace transformation to (1.55) gives

$$\hat{p}(x,y;\alpha) = \hat{p}(y,y;\alpha)\hat{F}(x,y;\alpha), \qquad x,y \in S,$$

and by replacing $\hat{p}(y,y;\alpha)$ by (1.57) we get for (1.58)

$$\hat{d}(x,y;\alpha) = \hat{d}(y,y;\alpha) + \frac{\pi(y)(\hat{F}(x,y;\alpha)-1)}{\alpha \hat{F}_\pi(y;\alpha)}, \qquad x,y \in S.$$

Regarding the limit $\alpha \to 0$ and applying l'Hospital's rule gives us

$$\begin{aligned} &\lim_{\alpha \to 0} \frac{\pi(y)(\hat{F}(x,y;\alpha)-1)}{\alpha \hat{F}_\pi(y;\alpha)} \\ &\quad = \lim_{\alpha \to 0} \frac{\pi(y)\int_0^\infty (-t)e^{-\alpha t} F(x,y;dt)}{\underbrace{\int_0^\infty e^{-\alpha t} F_\pi(y;dt)}_{\to 1} + \underbrace{\alpha \int_0^\infty (-t) e^{-\alpha t} F_\pi(y;dt)}_{\to 0}}, \qquad y \in S. \end{aligned}$$

Since $\lim_{\alpha \to 0} \hat{d}(x,y;\alpha) = d(x,y)$ follows from a Tauberian theorem we arrive at

$$d(x,y) = d(y,y) - \pi(y)m(x,y), \qquad x,y \in S,$$

and with $m(y,y) = 0$ which stems from our definition of the mean first passage time T_y the representation of the density of the deviation operator presented in Theorem 1.6.2 directly follows.

$\blacksquare$

Now we first of all need to compute $m(x,y)$ what we will exemplary do for the Brownian motion to derive a closed-form representation of the density of its deviation operator afterwards.

Let $\mathcal{X}$ be a Brownian motion with drift coefficient $\mu \in \mathbb{R}$ and diffusion parameter $\sigma^2 > 0$ both independent of $x \in S$. To justify our computations, we assume that the diffusion is reflected at the boundaries l and r and in case that $r = \infty$ we suppose that r is inaccessible (i.e., it cannot be reached within finite time); see [135]. We will start with the computation of the stationary distribution by (1.51) and (1.52) and get

$$\pi(x) = \frac{2\mu}{\sigma^2} \exp\left[\frac{2\mu}{\sigma^2}x\right] \left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1}, \qquad x \in S, \qquad (1.59)$$

(see, e.g., [57]) which exists whenever $\mu \cdot l < \infty$ and $\mu \cdot r < \infty$. Note that the existence of a unique stationary distribution is in accordance with our assumption of positive Harris recurrent and aperiodic processes. In case of

$S = [0, \infty)$ a negative drift $\mu < 0$ assures the existence of π and (1.59) simplifies to

$$\pi(x) = -\frac{2\mu}{\sigma^2} \exp\left[\frac{2\mu}{\sigma^2} x\right], \qquad x \in [0, \infty). \tag{1.60}$$

Now we will derive the mean first passage times. We have

$$m(x, y) = E[T_y | X_0 = x], \qquad x, y \in S,$$

as the mean first passage time to y starting in x with $m(x, x) = 0$, which reflects the fact that the average time needed to reach x when the process is already in x will obviously be equal to zero for all $x \in S$. We set

$$\acute{m}(x, y) = E[T_y | X_0 = x], \qquad x, y \in S,\ x < y,$$

and

$$\grave{m}(x, y) = E[T_y | X_0 = x], \qquad x, y \in S,\ x > y.$$

We focus on the computation of $\acute{m}(x, y)$ using a differential equation of second order. Basically, the computation of $\grave{m}(x, y)$ can be executed analogously and we will point out any differences in the derivation. Instead of starting the process in a fixed initial state $X_0 = x$, $x < y$, we condition the expectation on the arbitrary position of X_{t^*} with t^* sufficiently small so that the probability of reaching y within time interval $[0, t^*]$ is negligible, and get the random variable $\acute{m}(X_{t^*}, y)$ which we can represent as a Taylor series in x

$$\acute{m}(X_{t^*}, y) = \frac{(X_{t^*} - x)^0}{0!} \acute{m}(x, y) + \sum_{n=1}^{\infty} \frac{(X_{t^*} - x)^n}{n!} \frac{\partial^n \acute{m}(x, y)}{\partial x^n}. \tag{1.61}$$

Applying $E[\cdot | X_0 = x]$ to (1.61) gives

$$E[\acute{m}(X_{t^*}, y) | X_0 = x] = \acute{m}(x, y) + \sum_{n=1}^{\infty} \frac{1}{n!} E[(X_{t^*} - x)^n | X_0 = x] \frac{\partial^n \acute{m}(x, y)}{\partial x^n} \tag{1.62}$$

where

$$E[X_{t^*} - x | X_0 = x] = \mu t^* + o(t^*) \qquad \text{and} \qquad E[(X_{t^*} - x)^2 | X_0 = x] = \sigma^2 t^* + o(t^*)$$

(with $o(t^*)$ an error term and $\lim_{t^* \to 0} \frac{o(t^*)}{t^*} = 0$) follows directly from the fact that the increments of a Brownian motion $X_t - X_0$ are normally distributed with $\mathcal{N}(\mu t, \sigma^2 t)$. Furthermore, it holds

$$E[\acute{m}(X_{t^*}, y) | X_0 = x] = \acute{m}(x, y) - t^*.$$

Now we replace $E[\acute{m}(X_{t^*}, y)|X_0 = x]$ and $E[(X_{t^*} - x)^n|X_0 = x]$, $n \leq 2$, in (1.62) and obtain

$$\begin{aligned}\acute{m}(x,y) - t^* &= \acute{m}(x,y) + \mu t^* \frac{\partial \acute{m}}{\partial x} + \frac{1}{2}\sigma^2 t^* \frac{\partial^2 \acute{m}}{\partial x^2} + o(t^*) \\ &\quad + \sum_{n=3}^{\infty} \frac{1}{n!} E[(X_{t^*} - x)^n|X_0 = x] \frac{\partial^n \acute{m}}{\partial x^n}\end{aligned}$$

which we can transform to

$$-1 = \mu \frac{\partial \acute{m}}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 \acute{m}}{\partial x^2} + \frac{o(t^*)}{t^*} + \sum_{n=3}^{\infty} \frac{1}{n!} \frac{E[(X_{t^*} - x)^n|X_0 = x]}{t^*} \frac{\partial^n \acute{m}}{\partial x^n}.$$

Letting t^* tend to zero eliminates the error $o(t^*)$ and with

$$\lim_{t^* \to 0} \frac{E[(X_{t^*} - x)^n|X_0 = x]}{t^*} = 0, \qquad n \geq 3,$$

(see (1.4) in [73], p.160) we get

$$-1 = \mu \frac{\partial \acute{m}}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 \acute{m}}{\partial x^2}. \tag{1.63}$$

Now we have to compute the mean first passage time by solving the inhomogeneous second order differential equation (1.63). Transforming (1.63) gives us

$$-\frac{2}{\sigma^2} = \frac{2\mu}{\sigma^2} \frac{\partial \acute{m}}{\partial x} + \frac{\partial^2 \acute{m}}{\partial x^2} \tag{1.64}$$

with the solution $\acute{m}(x, y)$ being the sum of the general solution of the homogeneous equation $\acute{m}_h(x, y)$ and an individual solution $\acute{m}_i(x, y)$, $x < y$. For the first summand we get by using the auxiliary equation $\lambda^2 + \frac{2\mu}{\sigma^2}\lambda = 0$ with the two solutions $\lambda_1 = 0$ and $\lambda_2 = -\frac{2\mu}{\sigma^2}$

$$\acute{m}_h(x,y) = c_1 \exp[\lambda_1 x] + c_2 \exp[\lambda_2 x] = c_1 + c_2 \exp\left[-\frac{2\mu}{\sigma^2}x\right], \quad x < y,\ c_1, c_2 \in \mathbb{R}.$$

Furthermore, it can be easily seen that $\acute{m}_i(x, y) = -\frac{x}{\mu}$ is an individual solution to (1.64). Thus, we have

$$\acute{m}(x,y) = c_1 + c_2 \exp\left[-\frac{2\mu}{\sigma^2}x\right] - \frac{x}{\mu}, \qquad x < y,\ c_1, c_2 \in \mathbb{R},$$

where it remains to determine the two constants c_1, c_2. Applying $\lim_{x\uparrow y} \acute{m}(x,y) = 0$ (in case of $\grave{m}$ we have $\lim_{x\downarrow y} \grave{m}(x,y) = 0$) reduces the unknown coefficients to only one

$$0 = c_1 + c_2 \exp\left[-\frac{2\mu}{\sigma^2} y\right] - \frac{y}{\mu} \qquad \Leftrightarrow \qquad c_1 = \frac{y}{\mu} - c_2 \exp\left[-\frac{2\mu}{\sigma^2} y\right]$$

which leads to

$$\acute{m}(x,y) = \frac{y-x}{\mu} - c\left(\exp\left[-\frac{2\mu}{\sigma^2} y\right] - \exp\left[-\frac{2\mu}{\sigma^2} x\right]\right).$$

With the boundary condition for $x < y$ $\frac{\partial \acute{m}(x,y)}{\partial x}\Big|_l = 0$ we get

$$\left.\frac{\partial \acute{m}(x,y)}{\partial x}\right|_l = -\frac{1}{\mu} + c\left(-\frac{2\mu}{\sigma^2}\right)\exp\left[-\frac{2\mu}{\sigma^2} l\right] \qquad \Leftrightarrow \qquad c = -\frac{\sigma^2}{2\mu^2}\exp\left[\frac{2\mu}{\sigma^2} l\right]$$

and for $x > y$ $\frac{\partial \grave{m}(x,y)}{\partial x}\Big|_r = 0$ we get

$$\left.\frac{\partial \grave{m}(x,y)}{\partial x}\right|_r = -\frac{1}{\mu} + c\left(-\frac{2\mu}{\sigma^2}\right)\exp\left[-\frac{2\mu}{\sigma^2} r\right] \qquad \Leftrightarrow \qquad c = -\frac{\sigma^2}{2\mu^2}\exp\left[\frac{2\mu}{\sigma^2} r\right].$$

Hence, we have for the mean first passage times of a Brownian motion with $\mu, \sigma^2 \neq 0$ and $x, y \in S$

$$m(x,y) = \begin{cases} \frac{y-x}{\mu} + \frac{\sigma^2}{2\mu^2}\left(\exp\left[\frac{2\mu}{\sigma^2}(l-y)\right] - \exp\left[\frac{2\mu}{\sigma^2}(l-x)\right]\right) & x < y, \\ \frac{y-x}{\mu} + \frac{\sigma^2}{2\mu^2}\left(\exp\left[\frac{2\mu}{\sigma^2}(r-y)\right] - \exp\left[\frac{2\mu}{\sigma^2}(r-x)\right]\right) & x > y, \\ 0 & x = y, \end{cases} \tag{1.65}$$

which is the generalisation of the result in Problem 16 in [73], p.386.

Recall representation (1.53). We start with the computation of

$$\int_S \pi(z) m(z,y) dz, \qquad y \in S. \tag{1.66}$$

In case of a Brownian motion we get for (1.66) by inserting (1.59) and (1.65)

$$
\begin{aligned}
&\int_S \pi(z)m(z,y)dz \\
&\quad = \frac{2y}{\mu} - \frac{2}{\mu}\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1}\left(r\exp\left[\frac{2\mu}{\sigma^2}r\right] - l\exp\left[\frac{2\mu}{\sigma^2}l\right]\right) \\
&\quad +\frac{\sigma^2}{2\mu^2}\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1} \\
&\qquad \cdot\left(\exp\left[\frac{2\mu}{\sigma^2}(2r-y)\right] - \exp\left[\frac{2\mu}{\sigma^2}(2l-y)\right]\right). \qquad (1.67)
\end{aligned}
$$

In case of $S = [0,\infty)$ (1.67) simplifies to

$$
\int_S \pi(z)m(z,y)dz = \frac{2y}{\mu} + \frac{\sigma^2}{2\mu^2}\exp\left[-\frac{2\mu}{\sigma^2}y\right].
$$

Now we can compute the density of the deviation operator of a Brownian motion by inserting (1.59) and (1.67) into (1.53) and get for $x = y$

$$
\begin{aligned}
d(y,y) &= \pi(y)\int_S \pi(z)m(z,y)dz \\
&= \frac{4y}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right]\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1} \\
&\quad -\frac{4}{\sigma^2}\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-2} \\
&\qquad \cdot\left(r\exp\left[\frac{2\mu}{\sigma^2}(r+y)\right] - l\exp\left[\frac{2\mu}{\sigma^2}(l+y)\right]\right) \\
&\quad +\frac{1}{\mu}\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-2}\underbrace{\left(\exp\left[\frac{4\mu}{\sigma^2}r\right] - \exp\left[\frac{4\mu}{\sigma^2}l\right]\right)}_{\left(\exp\left[\frac{2\mu}{\sigma^2}r\right]\right)^2 - \left(\exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^2}
\end{aligned}
$$

what we transform to

$$\begin{aligned} d(y,y) &= \left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1} \\ &\qquad \cdot\left(\frac{4y}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}r\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}l\right]\right) \\ &\qquad -\frac{4}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right]\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-2} \\ &\qquad \cdot\left(r\exp\left[\frac{2\mu}{\sigma^2}r\right] - l\exp\left[\frac{2\mu}{\sigma^2}l\right]\right). \end{aligned}$$

For $x < y$ we obtain

$$\begin{aligned} d(x,y) &= d(y,y) - \pi(y)m(x,y) \\ &= \left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1} \\ &\quad \cdot\left(\frac{2(x+y)}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}r\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}(l+y-x)\right]\right) \\ &\quad -\frac{4}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right]\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-2} \\ &\quad \cdot\left(r\exp\left[\frac{2\mu}{\sigma^2}r\right] - l\exp\left[\frac{2\mu}{\sigma^2}l\right]\right) \end{aligned}$$

and for $x \geq y$

$$\begin{aligned} d(x,y) &= \left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-1} \\ &\quad \cdot\left(\frac{2(x+y)}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}l\right] + \frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}(r+y-x)\right]\right) \\ &\quad -\frac{4}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right]\left(\exp\left[\frac{2\mu}{\sigma^2}r\right] - \exp\left[\frac{2\mu}{\sigma^2}l\right]\right)^{-2} \\ &\quad \cdot\left(r\exp\left[\frac{2\mu}{\sigma^2}r\right] - l\exp\left[\frac{2\mu}{\sigma^2}l\right]\right). \end{aligned}$$

For the density of the deviation operator of a reflected Brownian motion $\mathcal{X}$ on $S = [0,\infty)$ with $\mu < 0$ it holds

$$d(x,y) = -\frac{1}{\mu}\exp\left[\frac{2\mu}{\sigma^2}(y - \min\{x,y\})\right] - \frac{2(x+y)}{\sigma^2}\exp\left[\frac{2\mu}{\sigma^2}y\right], \quad x,y \in [0,\infty).$$

Naturally, it is also possible to derive $d(x, y)$, $x, y \in S$ for more general diffusion processes. However, this requires the solution of the inhomogeneous second order differential equation (1.63) with variable coefficients $\mu(x)$ and $\sigma^2(x)$, $x \in S$, and is a topic of future research.